STATISTICAL MODELING OF RADIATION DETECTION AND IMAGING SYSTEM

BY

NAN LI

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Nuclear Engineering
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2012

Urbana, Illinois

Doctoral Committee:

Associate Professor Ling-Jian Meng, Chair, Director of Research
Professor James Stubbins
Professor Rizwan Uddin
Professor Zhi-Pei Liang

Abstract

In the development of SPECT system for imaging small animal *in vivo*, higher spatial resolution is continuously sought to provide improved spatial detail and image quality. However, the pursuit of high resolution often results in the poor photon collection efficiency and limited quantitative accuracy, which ultimately confines the capacities of the imaging modality. Adaptive imaging approach could be implemented in order to maximize the efficiency for collecting useful imaging information regarding a given task and therefore provide an optimum image performance. In the adaptive data acquisition, the system hardware or acquisition protocol could vary in response to information being acquired during an imaging study, which requires a reasonable optimization method to conduct the system configuration alteration.

In practical system optimization process, the key challenge is to optimize the system performance in real-time with respect to a wide range of design and imaging parameters for observing the unknown object. Even with state-of-the-art parallel computing platform, we still have to face two main difficulties: large computation load for performance indices evaluation and the complexity of the optimization problem against different system parameters. In order to address these issues, we have developed a series of approaches to enable SPECT system optimization with reasonable computation load.

As the first step towards the system optimization, we have developed a vector modified uniform Cramer-Rao bound (MUCRB) to replace the time-consuming brute-force MC simulation. This approach allows one to analytically derive the fundamental tradeoffs between resolution and minimum achievable total (or average) variance over arbitrarily chosen voxels, which could be asymptotically achieved by the post-filtered penalized maximum likelihood estimation with well-defined penalty function in linear Poisson model. In order to further reduce the computation load, the non-uniform object-space pixelation (NUOP) approach has been developed to divide the object-space into smaller voxels for target-regions, and into larger voxels in areas that are relatively smooth and/or less important to the reconstruction of the target-regions. This method could improve the calculation speed for reconstruction and vector MUCRB calculation by 1-2 orders of magnitude without sacrificing image quality inside the target-regions. The combination of these two approaches could allow real-time system performance evaluation.

Based on these techniques, we have also developed a generic system optimization method that allows the system to be optimized against any arbitrarily given system parameters. This approach was first used to develop an adaptive angular sampling approach for SPECT imaging. This

approach allows the camera to spend larger fractions of imaging time at angles those are relatively more efficient for acquiring useful imaging information from the target to deliver significantly lowered image variance at a given resolution, which builds a foundation for further optimizing performance with respect to many other system parameters. We have further expanded this method to develop an indirect system optimization approach dealing with a discretized the parameter space. It helps to identify the best combination of different system configurations to be used for attaining the optimum imaging performance. The series of analytical approaches developed through this thesis work could be used together to provide an efficient computation scheme to facilitate real-time system optimization. It could be used along with variable hardware detection system for use in adaptive SPECT imaging.

To Nathan, Kevin, Cong
and
Mom and Dad
who help to see what's important and what's not

ACKNOWLEDGMENTS

First of all, I would like to express profound gratitude to my advisor, Professor Ling Jian Meng, for giving me his guidance, support and patience during the course of the research. Moreover, I am inspired by his insights into the evolution of high resolution SPECT imaging, his attention to detail and his dedication to the profession. I am also grateful to my committee member, Professor James Stubbins, for his helpful feedback during the preparation of my dissertation.

I would also like to thank all of my co-workers in the radiation imaging group. I am appreciative to Geng Fu, Jiawei Tan, and Liang Cai for insightful discussions and for their assistances in experimental hardware installation for implementing and demonstrating the system performance evaluation and optimization methods I've developed.

In addition, I am indebted to my parents and my parents-in-law for their trust and for their incredibly support during my graduate life. They taught me the value of the family.

Lastly and most importantly, I would like to give special thanks to my wife, Cong Zhou, who endured my many hours away from family during the course of this Ph.D. program, and still gave me the love, the support I needed and the patience I often did not deserve. Her unconditional love and enthusiastic spirit have enabled me to become the person I am today. Furthermore, the priceless gifts I receive from her are our sons, Nathan and Kevin, who usually remind me not to work too hard and help me to find out the new fun for life I have never experienced before.

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION TO EMISSION TOMOGRAPHY

In recent years, emission tomography becomes a widely-used imaging approach to study the biological function in humans and animals *in vivo* [1]. As the term *emission tomography* suggests, this technique produces a three-dimensional images of the internal structures using gamma-ray emission. In comparison to X-ray computed tomography (CT) that principally depicts the body's architectural structure, emission tomography is categorized as a functional imaging approach, which can measure the spatial distribution of radioactive materials inside the tissue to detect abnormalities that are characteristic of diseases, such as diagnose of tumors, and/or to image properties of the body's physiology, such as shape of glucose uptake affected by Parkinson's disease, growth rate of activities in tumors influenced by drugs and so on.

The emission tomography encompasses two main imaging modalities: positron emission tomography (PET) and single photon emission computed tomography (SPECT). Even with the dramatic development of imaging methods, we still face some challenges for pursuing outstanding image quality and system performance.

1.1 Positron Emission Tomography

Positron emission tomography (PET) is a nuclear medicine imaging approach that relies on the external detection of annihilation gamma rays administered to the tracer distribution in the target object. In PET imaging, certain positron-emitting radionuclide tracers are introduced into the body on biologically active molecules, and positrons could be generated by their radioactive decays. As one positron generally travels through human tissue, it will give up its kinetic energy to very low, near the end of its range, by Coulomb interactions with negative electrons. Then the positron could combine with a normal electron by the formation of a hydrogen-like orbiting pair called positronium, which is unstable and eventually decays in the process of annihilation. So the original positron and electron disappear and are replaced by a pair of anti-parallel 511keV photons [2]. The PET detector system often consists of a ring of detectors to collect pairs of oppositely directed gamma rays simultaneously, which could be processed to reconstruct a three-dimensional dataset by some computing algorithms [3]. This dataset may then be manipulated to show thin cross-sectional slices along any chosen axis of the body [4].The PET imaging process

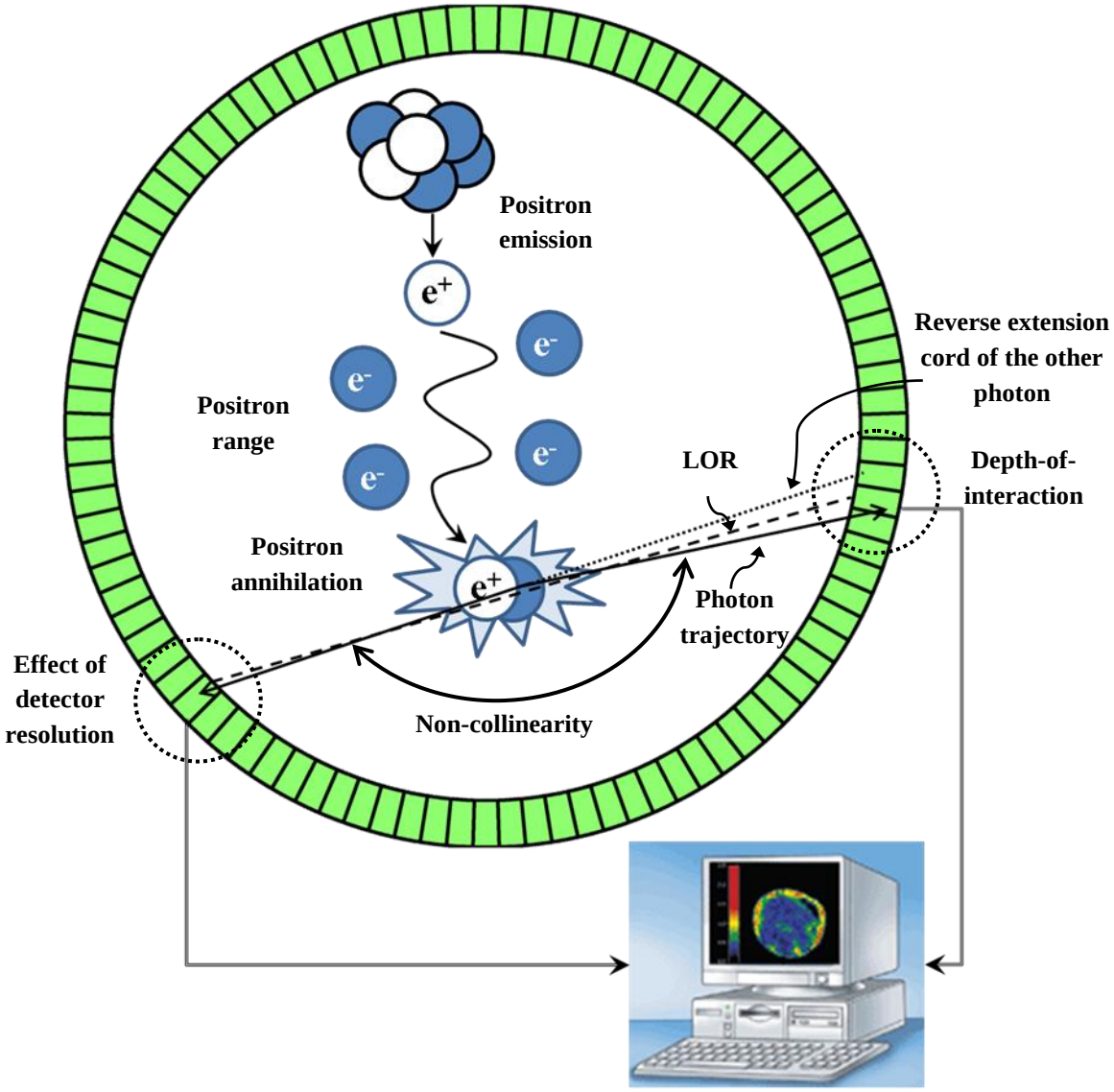

Figure 1.1 Physics of positron decay and annihilation, with depiction of the finite positron range, the non-collinearity of the annihilation photons, effect of detector resolution and depth-of-interaction inherent to the positron–electron annihilation process which give rise to a fundamental positional inaccuracy. Positron range and angle are to a great extent exaggerated.

is shown in Figure 1.1 and an example of a PET reconstructed image of human brain is given in Figure 1.2 [5].

The common radioisotopes used in PET are ^{18}F, ^{11}C, ^{13}N, ^{15}O, and ^{82}Rb. The properties of these isotopes are given in Table 1.1. On the one hand, ^{11}C, ^{13}N, and ^{15}O have very short half-lives of just more than one minute. The practical application of pharmaceuticals labeled with these isotopes requires that cyclotron and radiochemistry facilities are built on-site for local production. On the other hand, ^{18}F-labelled compounds, with a half-life of almost 2 hours, could greatly

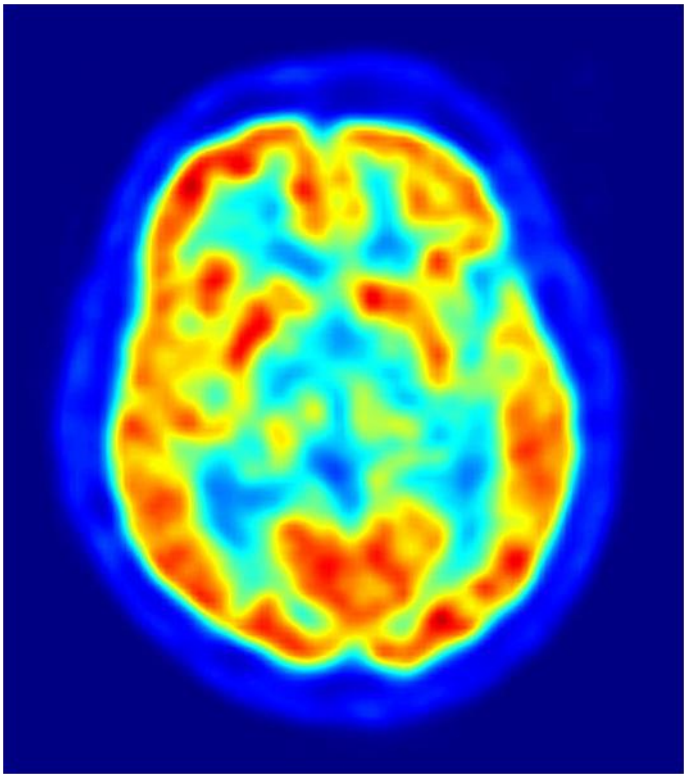

Figure 1.2 Transaxial slice of the brain of a 56 year old male patient taken with PET

facilitate the use of PET imaging. Moreover, although ^{82}Rb has a very short half-life as well, it can be distributed as it can be produced via a ^{82}Sr generator system, which has a half-life of 25 days and a practical lifetime of 1 to 2 months [6].

Table 1.1 Properties of Common PET Radioisotopes

Isotope	Half-life (min)	Maximum Positron Energy (MeV)	Positron Range in Water (mm)[1]
^{11}C	20.3	0.96	2.1
^{13}N	9.97	1.19	0.57
^{15}O	2.03	1.7	1.02
^{18}F	109.8	0.64	1.2
^{82}Rb	1.26	3.15	12.4

1. The range is defined as the radius that includes 75% of all annihilation events

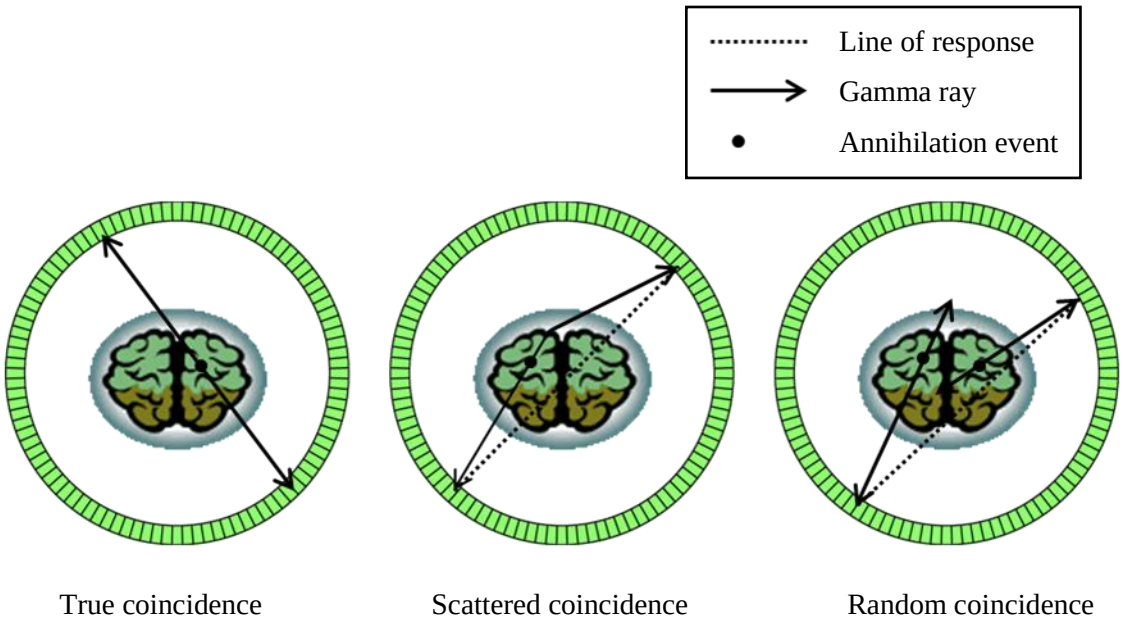

Figure 1.3 the three types of coincidence events measured in a PET system

As seen in Figure 1.1, localization of an annihilation event could be determined by a line of response (LOR) across the object joining the two relevant detectors. It is desired to obtain excellent image quality by measuring a great number of photons for good statistical precision and acquiring sufficient information carried by detected photons for accurate localization determination and thus high spatial resolution.

A substantial advantage of PET imaging over SPECT is a much higher sensitive (by approximately two to three orders of magnitude). A PET system typically consists of single ring or multiple rings of detectors that are working in coincidence. This design enables the patient to be completely surrounded by detectors. The absence of mechanical collimation ensures high collection efficiency to annihilation photons and thus the minimization of radiation dose. As shown in Figure 1.3, measurement in PET imaging undergoes the mixture of true coincidence, scattered coincidence and random coincidence. A scattered coincidence is that one or both photons from a single event are scattered and both are detected. The changing direction of the photon is likely to result in the wrong LOR assigning the coincidence event. Since their energies are lost, scattered events could be rejected by an excellent energy window setting around the 511 KeV. A random coincidence occurs when two separate photons do not arise from the same annihilation event but are close enough in time to be incidentally detected as a coincidence. A narrow coincidence time window setting could be utilized during data acquisition to reduce random events without compromising true coincidence sensitivity. Therefore, the system

sensitivity could be increased by measuring as many true coincidence events as possible and minimizing the scattered and random coincidences.

In addition, solid angle coverage of the detectors could also affect the sensitivity. In general, the total solid angle can be improved by either increasing number of detectors or pushing detector closer to the object. At last, one could obtain excellent system sensitivity using a thick detector with high atomic number and density. For a PET imaging system, high sensitivity enables high statistical performance of acquired data and help to achieve an outstanding image quality.

Even with the excellent sensitivity, however, spatial resolution performance in PET imaging could be degraded by a number of different factors: positron range, photon non-collinearity, intrinsic spatial resolution of the detector and effect of depth of interaction in the detector [1].

At first, the spatial resolution of PET imaging is limited by the fundamental nature of positron annihilation. Once emitted, the positron deposits its kinetic energy in the process of Coulomb interaction with electrons in the tissue. Since the rest mass of positron is the same as that of electron, positron movement could suffer from large deviations in direction, seen in Figure 1.1. The positron range could be defined as the average distance of emitted positrons traveling in the surrounding medium before they can reach thermal energies in order to be annihilated [7]. As a result, the uncertainty of positron range could bring ambiguity in determining the location of radioactive decay. Different positron-emitting isotopes reveal distinct energy distributions, and therefore exhibit different positron range values in the statistical sense, seen in Table 1.1.

Traditionally, positron range has been viewed as a purely resolution-limiting factor. However, one could attempt to reduce this effect by two additional approaches (one hardware-based and the other software-based). On the one hand, the positron range is known to be reduced in a strong magnetic field. It must be noted that this effect is significant for high-energy positron-emitting radionuclide, e.g., ^{82}Rb, at magnetic field of 5T or more. This is one of the advantages of designing MRI-compatible PET imaging system [8-10]. On the other hand, even though it is impossible to determine the positron range for each particular detected event, one could calculate and incorporate the related probability distribution into the system response matrix for statistical reconstruction algorithms. Such advanced modeling in turn can result in improvements in image resolution [3, 11, 12].

In addition to the positron range, non-collinearity also leads to a limitation of the spatial resolution of PET imaging. Since the net momentum for an emitted positron and the related electron with which it annihilates can be non-zero, there exists an angular uncertainty in the trajectories of the 511 KeV photons due to conservation of momentum. This deviation is approximately 0.25° [13, 14]. As a result, the annihilation gamma rays are not exactly anti-

parallel, shown in Figure 1.1, which is referred to as non-collinearity. The corresponding resolution blurring depends on the detector ring diameter *D*, and is approximately given by:

$$FWHM \approx (0.25 \times \frac{\pi}{180}) \frac{D}{2} \approx 0.0022 \times D. \quad (1.1)$$

Therefore, one could expect about 1.54 mm ~ 1.76 mm FWHM blurring for a typical whole-body scanner (*D* is about 70 ~ 80 cm), and only 0.17mm FWHM blurring for animal scanners (*D* is roughly 8 cm) [1].

The third significant factor limiting PET image resolution is the intrinsic spatial resolution of the detector. For a detector composed of small discrete crystals, all interactions are assumed to occur at the center of individual crystals for the purpose of backprojection and image reconstruction, illustrated in Figure 1.1. As a result, the finite spatial resolution of detector crystals must introduce uncertainty in determining the locations of photon detections [15]. In general, the cost consideration and technical developments could limit the detector intrinsic resolution.

A final factor affecting PET imaging resolution is the depth of interaction of the gamma rays in the crystal, seen in Figure 1.1. In photon detection, a gamma ray travels in the crystal before being completely absorbed. If the gamma ray enters the crystal at an oblique angle, the location of the interaction will not be the same as the point of entry into the crystal. Thus, unless the depth of interaction within a crystal can be accurately determined, an incorrect line of response, assigned to a position at the front of the crystal of interaction, will lead to ambiguity in reconstructed images [16-18].

In recent years, some new detectors and techniques have been rapidly developed [19-25]. For example, semiconductor detector crystals, e.g. CZT or cadmium telluride, could be employed instead of scintillation crystals. The high atomic number and high density gives this detector satisfied sensitivity performance. Furthermore, this detector is able to achieve high-resolution 3-D positioning and superior energy resolution. In order to make optimal use of these excited developments and obtain the optimum system sensitivity and spatial resolution performance in reconstructed images, it is desired to optimize PET system design with respect to many system parameters, such as detector ring diameter, crystal transaxial width, crystal radial length and so on. As introduced above, some parameters are inextricably intertwined, and even counteractive to each other [26-28]. For example, thinner thickness of detector could alleviate the effect of depth of interaction, at the cost of the reduced detection efficiency. Decreasing the ring diameter of PET detectors could lead to the mitigation of photon non-collinearity, but the depth of interaction may become serious since a larger fraction of the gamma rays could enter detectors at oblique angles.

Therefore, a practical system performance optimization method is necessary for optimum image quality in PET imaging.

1.2 Single Photon Emission Computed Tomography

Single photon emission computed tomography (SPECT) is another diagnostic imaging technique in which the images of radionuclide distribution are generated *in vivo* based on the measurements of gamma rays as well. Different from positron emission tomography (PET), however, these photons are directly generated from the site of the radioactive decays in the object. In SPECT imaging, target-specific biological radiolabeled tracers are introduced to probe molecular processes associated with the disease being assessed both at the cellular and sub-cellular levels within living organisms. Compared with PET isotopes which are positron emitters, SPECT isotopes own many advantages. At first, there exists a wide variety of radiotracers applicable to SPECT imaging, which can be cheaper, easier to acquire, and, in certain tumors, more accurate. Several available isotopes are listed in Table 1.2. In addition, the radiotracers used in SPECT have relatively long half-lives from a few hours to a few days, such as ^{99m}Tc, ^{111}In, ^{123}I and ^{201}Tl, seen in Table 1.2, which can be tailored to some specific applications. Furthermore, SPECT imaging is able to image two or more radiotracers simultaneously and then delineate the correlations between different biological or molecular targets. As a result, the SPECT dual-tracer imaging can provide additional diagnostic value that is difficult for PET single-tracer imaging. This method is mainly based on the separation of the gamma photons by applying different energy windows for different gamma photons [29-32].

Table 1.2 Radio Isotopes Useful in Medical Imaging Technology

Isotope	Half-life	Gamma Energies (KeV)
^{18}F	109.8 minutes	511
^{99m}Tc	6.0 hours	140
^{111}In	2.8 days	245, 172
^{123}I	13.2 hours	159, 529
^{125}I	59.6 days	27.4, 35.5
^{11}C	20.4 minutes	960
^{201}Tl	3.0 days	167, 135

SPECT imaging usually struggles with tradeoffs between resolution property and statistical noise level. Higher spatial resolution is continuously sought in nuclear imaging system to

improve the spatial detail and image quality which could be used for preclinical and clinical imaging studies. Because it is not necessary to consider physics-related limitations of positron range and photon collinearity involved in PET imaging, SPECT is potentially capable of providing much more excellent spatial resolution over PET. However, the use of collimation could lead to a well-known bottleneck for high resolution SPECT instrumentations: the limited sensitivity and thus relatively poor statistical property.

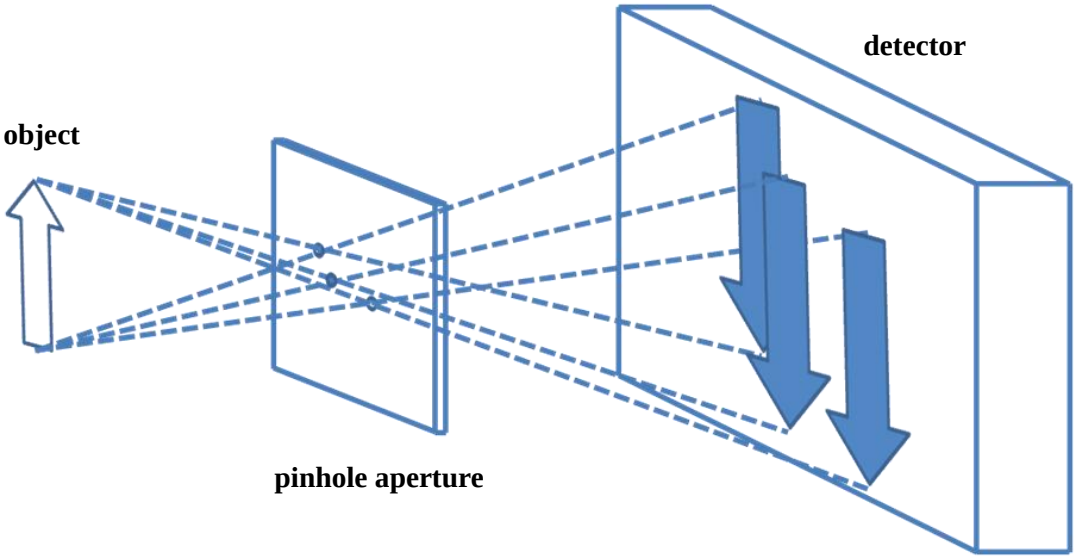

Figure 1.4 SPECT imaging model using a multi-pinhole aperture

Small pinhole aperture with large magnification is accepted as the standard method for considerable enhancing spatial resolution in SPECT, especially in small animal imaging, which could be simply modeled in Figure 1.4. Small pinhole could mechanically select the possible detected photons traveling in specific directions with precision and others are shielded. Large magnification, limited by the size of Field-Of-View (FOV), could contribute to clear discrimination of individual detected photons in discrete detector space. As a result, one could obtain high information per detection to identify the origin of the detected photon, which could decrease blurs in reconstructed images and lead to excellent resolution property. However, the resulting small open-fraction of the aperture could also greatly reduce the total amount of detected photons within certain imaging time. The use of multi-pinhole could significantly increase the system sensitivity, but the success is limited by multiplexity, especially for large magnification. In this case, projections from different pinholes may be overlapped in the detector plane, seen in Figure 1.4. Obviously it would bring ambiguity about which pinhole captures detected photons, and thus reduce the information the photons carried for determining their origins. It was demonstrated in [33] that the gains in improved collection efficiency are offset by

increased multiplexing for image quality. As a result, although large magnification is beneficial to spatial resolution, it could restrict the increase of pinhole number and thus system sensitivity. In one word, there exists a tradeoff between information per detected photon and total amount of detected photons, which is usually reflected by balance between resolution and statistical noise.

In recent years several state-of-art small animal SPECT systems have been designed and constructed for good tradeoffs between spatial resolution and system sensitivity [33-34]. At first, let's take an ultra-high resolution single photon emission microscope (SPEM) system as an example, which has been developed in our lab for small animal imaging [40-42]. As shown in Figure 1.5, the SPECT imaging instrument consists of two independent imaging devices based on high spatial resolution columnar scintillators, in combination with image intensifiers and high sensitivity, high resolution Electron Multiplying Charge-Coupled Device (EMCCD) cameras, which could be modeled with 1024×1024 pixels of $48\,\mu m \times 48\,\mu m$ in size. During the image acquisition procedure, the animal is placed in a vertical holder which rotates in front of the imaging detector for sufficient angular sampling information. In order to guarantee both high spatial resolution and high sensitivity, the imaging devices use 7- and 19-pinhole collimators with 300 μm and 450 μm in diameter, respectively, which are placed between the detectors and objects to select the gamma rays from certain incident direction. During data acquisition, detectors could record multiple 2-D projections, which are then processed in tomographic reconstruction algorithms [45, 46], in parallel in principle, to yield 3-D images of tracer concentration within the body. Representative images of mice brain acquired in the SPEM system are shown in Figure 1.6. Because of the excellent performance of detector intrinsic resolution, the SPEM system could deliver an ultra-high imaging resolution of around 100 μm in phantom studies. It was also demonstrated that the current dual-headed SPEM system is capable of visualizing a very small number of radiolabeled cells in mouse brain. However, even with a large number of pinholes in the entire system, the overall efficiency for the SPEM system is typically 10^{-4} or lower, which ultimately confines the capacities of the system.

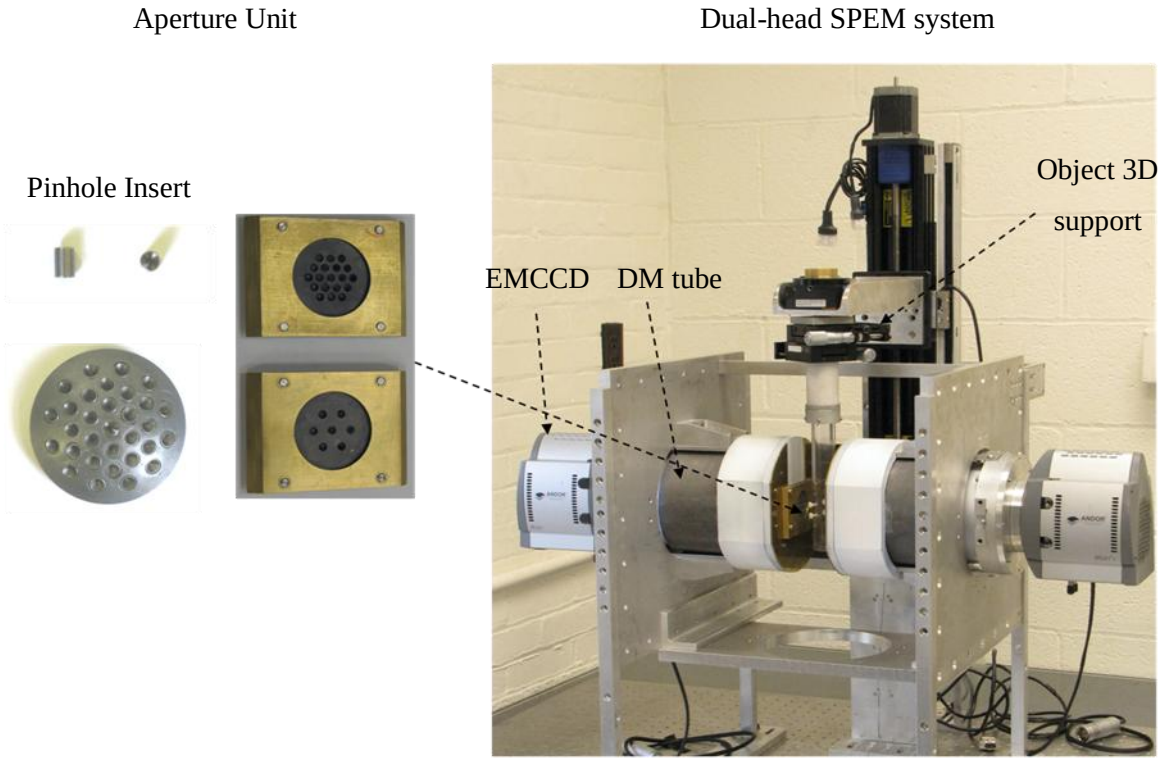

Figure 1.5 Ultra-high Resolution Single Photon Emission Microscope (SPEM) system

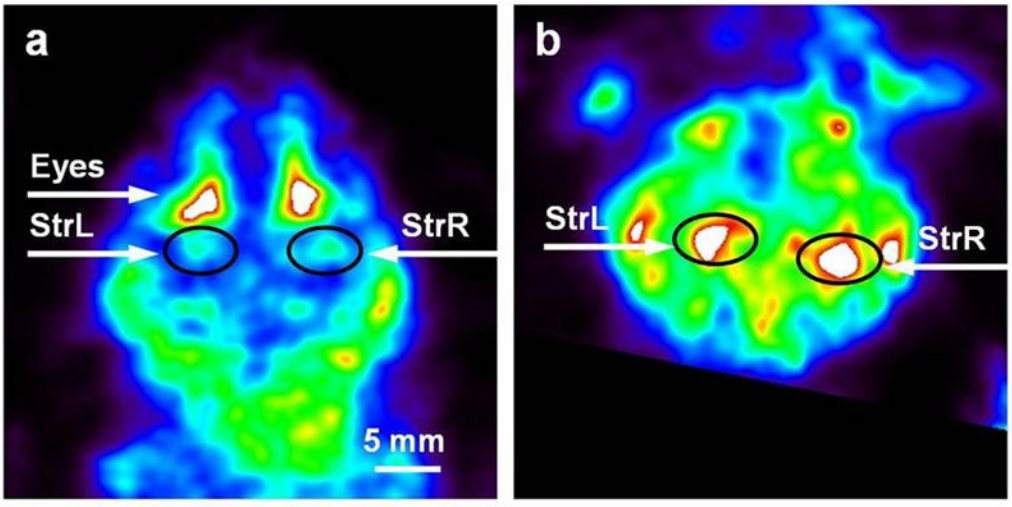

Figure 1.6 Mouse brain SPECT imaging. Intensity slices through the volumetric reconstruction of the mouse brain, labeled with [^{99m}Tc]TRODAT-1: A) transversal brain image; (B) coronal brain image at the level of the Striatum. StrL – Left Striatum; StrR – Rigth Striatum

In our lab we have also developed another high performance dual-head SPECT system, seen in Figure 1.7, which is based on a novel energy-resolved photon-counting (ERPC) detector [36]. In this system, the ERPC detector consists of eight CdZnTe or CdTe detector hybrids and

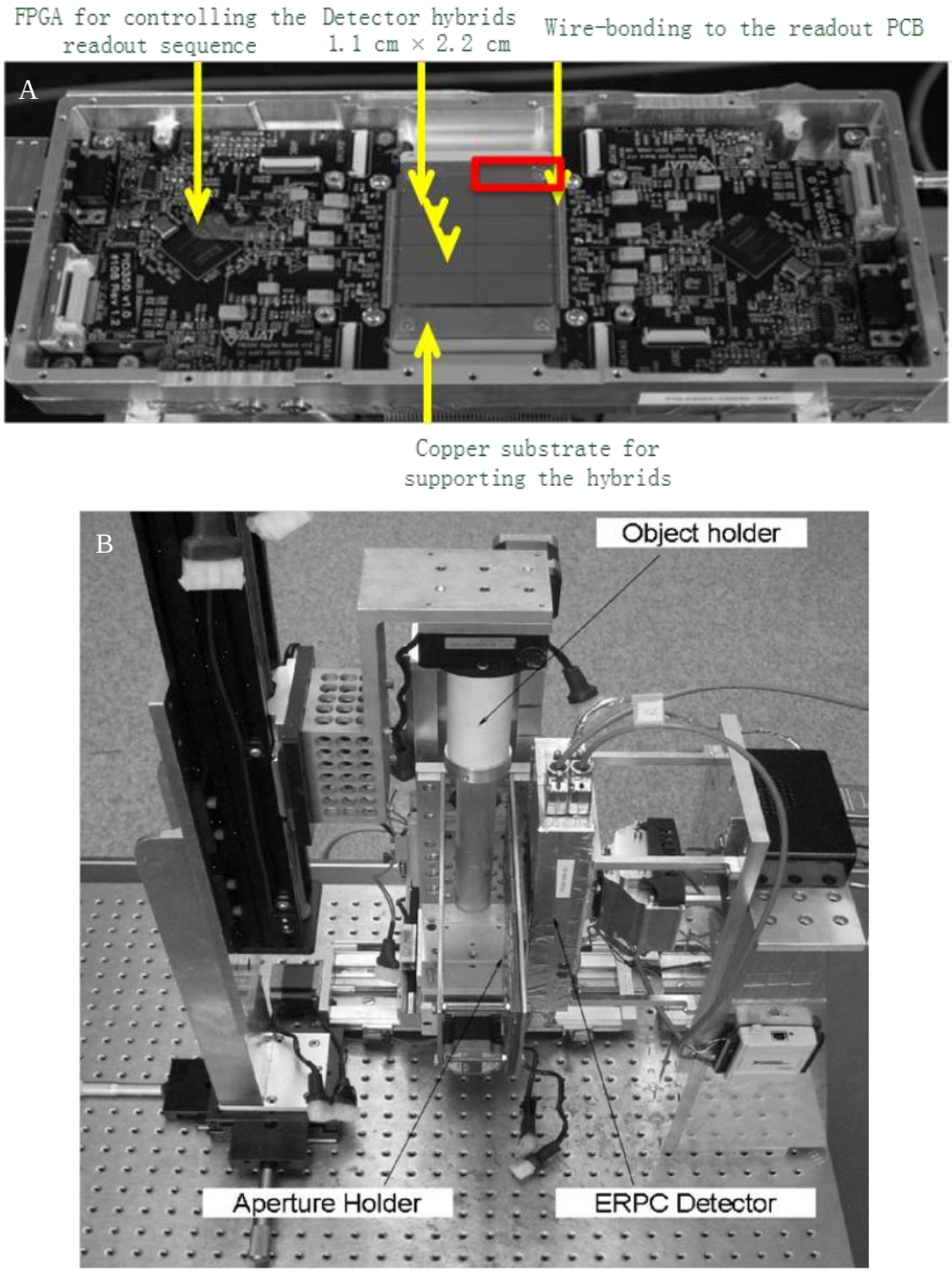

Figure 1.7 (A) The ERPC detector with eight detector hybrids. (B) General energy-resolved photon-counting (ERPC) detector platform.

provides an overall detection area of 4.5 cm × 4.5 cm. Each hybrid has an active area of approximately 11 mm × 22 mm and accommodates a total of 2048 readout channels arranged in an array of 32 × 64 pixels. These semiconductor detectors offer a combination of good energy resolution of around 3~4 keV at 140 keV, good intrinsic imaging resolution of 350 μm × 350 μm,

and adequate detection efficiency for gamma rays emitted by common single photon emitters, such as ^{125}I, ^{99m}Tc, ^{123}I and ^{201}Tl. ERPC detectors with both 1mm and 2mm thickness have both been developed. In order to fully utilize the excellent intrinsic detector resolution, each detector head is coupled to a multiple pinhole aperture that has 25 pinholes arranged in 5×5 pattern. The pinhole size is 200 μm in diameter. In preliminary study, the spatial resolution offered by the SPECT system should be less than 250 μm. Furthermore, this detector platform is capable of providing accurate estimation of 3D interaction position for each event to reject the effect of depth of interaction, which offers richer information per detected photon and thus equivalently reduces the requirement for the sensitivity.

In Netherland, Beekman et al. have developed another ultra-high resolution static SPECT system, called USPECT-II [34]. This system makes use of 3 ultra-large stationary NaI(Tl) broadband detector arrays (595mm x 472mm each) with fully digital electronics. A typical 10 times magnification factor for mice further contributes to the outstanding system resolution, which could reach less than 350 μm in 3D phantom study. Moreover, three optical cameras are placed outside of SPECT system to select the target area to be imaged, which will be moved to the focus of a 75-pinhole cylindrical collimator. The USPECT-II could obtain peak geometric sensitivity of 0.07% with 350 μm pinholes and of 0.18% with 600 μm pinholes.

The last example of SPECT system is the FastSPECT-III, developed by Miller et al. in the university of Arizona [37]. The stationary SPECT system is constructed using a new class of high-resolution BazookaSPECT gamma-ray cameras, which are the combination of a scintillation crystal, an integrating CCD/CMOS sensor, and a microchannel plate (MCP)-based image intensifier. This detector offers effective detector pixel size of 104 μm × 104 μm and high counting rate operation capability of more than 10^7 counts per second. In the system, as many as twenty independent BazookaSPECT detectors work together for high collection efficiency about a spherical FOV, each of which is coupled with a 200 μm single-pinhole aperture to be compatible with outstanding detector intrinsic resolution. As a result, FastSPECT-III system could provide spatial resolution of about 250 μm and satisfied system sensitivity.

For high resolution SPECT system, the development of the detector technology provides strong support for the pursuit of outstanding spatial resolution in images. In order to fully utilize the advanced detectors, the small pinhole size must be used, which restricts the performance of SPECT system by poor collection efficiency and limited quantitative accuracy. In order to maximize collection efficiency and obtain a superior balance between resolution and statistical noise level, many design and imaging parameters could be adjusted, e.g., the number and the size of the pinholes in apertures, locations of apertures in the system, magnifications, imaging time

distributions and so on. These parameters have complex influences on system performance, which requires a practical method to find an optimum set of configuration.

1.3 A Key Challenge to High Resolution Nuclear Imaging System

As previously discussed, there is a tradeoff between information per detected photon and total amount of detected photons in SPECT imaging, which could be represented by the balance between spatial resolution and system sensitivity in practice. In PET imaging, it is also desired to achieve high collection efficiency with excellent resolutions performance.

In recent years, developments of advanced detectors and novel techniques provide solid foundation to build a high resolution nuclear imaging system. Therefore, it is required in system design to maximize the efficiency for collecting useful imaging information regarding a given task, and therefore obtain an optimum image quality. The requirement could be satisfied by an adaptive imaging approach, reported by Barrett et al. [47], Clarkson et al. [48], Freed et al. [49] and Li et al. [50]. The adaptive imaging system could change its configuration or protocol during data acquisition in response to the acquired preliminary information in an imaging study. As a result, the adaptive imaging system could adapt itself to particular imaging task and offer improved collection efficiency.

Practical adaptive data acquisition in the case of real system is difficult for two reasons. At first, mechanical control of the imaging configuration is slower and more complex, which results in the practical difficulty of adaptive-mode implementation. This problem could be partially alleviated with the development of high-precision and extremely-compact equipments. Several adaptive system prototypes have been designed and constructed, such as prototype adaptive SPECT systems built by Freed et al. [49] and Cai et al. [51], a zoom-in adaptive PET geometry proposed by Zhou et al. [52] and a flexible x-ray micro-CT system developed by Moore et al. [53]. Alternatively, it may be very complicated to conduct the system configuration alteration in the adaptive imaging procedure. The key challenge for ultra-high resolution nuclear imaging system design is to optimize the system configuration with respect to a wide range of design and imaging parameters, especially for observing the unknown object. One potential problem in optimization application is that the optimum system configurations are required to be derived in a short time. In addition, system parameters are not independent to each other, but take the sophisticated influences together on the resolution and sensitivity. For example, In SPECT

imaging, the system sensitivity is controlled by the pinhole diameter and distance between the object and the aperture plane. Spatial resolution, which depends on pinhole diameter, detector resolution and magnification, can be traded off for sensitivity and field-of-view. Therefore, a practical optimization method is necessary as a meaningful adaptation rule in realization of adaptive imaging.

In order to address these issues, we will introduce four approaches developed for system performance optimization, which could be incorporated with many nuclear imaging systems equipped with variable hardware to realize the adaptive imaging and thus obtain the optimum image quality. In this work, SPECT is used as a general platform to demonstrate the feasibilities of the methods developed. It is worth noting that the developed optimization scheme could be applied to other high resolution imaging modalities, such as PET or CT.

1.4 Outline of the Dissertation

The remaining chapters of this dissertation are organized as follows:

Chapter 2 introduces some basic knowledge on system optimization. The system performance is usually quantified by resolution-variance tradeoffs for single voxel estimation and bias-variance tradeoffs for ROI quantification in our study. In practical system optimization process, we have to face two main difficulties: large computation load for performance indices evaluation and the complexity of the optimization problem against different system parameters. In order to address these issues, we have developed a series of approaches to enable SPECT system optimization with reasonable computation load, which will be detailed introduced one by one in the chapters later.

Chapter 3 reviews the development of the Cramer-Rao type bound. Based on it we have developed the vector modified Uniform Cramer-Rao bound (MUCRB) to analytically derive the fundamental tradeoffs between resolution and minimum achievable total (or average) variance over an arbitrary set of voxels offered by a given system design to avoid time-consuming Monte Carlo simulation. In addition, the vector MUCRB could also estimate the tradeoffs between the mean bias and the corresponding variance for region-of-interest (ROI) quantification. These tradeoffs predicted by vector MUCRB can be asymptotically achieved by post-filtered penalized maximum likelihood estimators with well-defined penalty function. The achievability of the MUCRBs ensures that the predicted system performance is meaningful in routine practice.

Chapter 4 presents the non-uniform object-space pixelation (NUOP) approach to adaptively divide the object-space with non-uniform pixel sizes according to the characteristics of the object and the input from the user. This approach could reduce the calculation load by 1-2 orders of

magnitude without scarifying image quality inside the ROI. Furthermore, the NUOP approach could be adapted to the vector modified Uniform Cramer-Rao bound for evaluating the image quality in the reconstruction in almost real time.

Chapter 5 describes an analytical adaptive angular sampling approach for efficiently optimizing the angular sampling strategy for a rotating SPECT system, which builds a foundation for further optimization with respect to many other parameters. This method allows the camera to spend larger fractions of imaging time at angles those are relatively more important for acquiring useful imaging information regarding a given imaging task. As a result, one could obtain greatly improved information collection efficiency and thus a significantly lowered image variance.

Chapter 6 proposed a general strategy for optimizing SPECT imaging system design or sampling strategy against a wide range of parameters. Instead of directly optimizing performance, this approach finds the optimum configuration by discretizing the parameter space and then optimizing the performance indices against imaging times assigned to individual possible system configurations using adaptive angular sampling approach. It could be demonstrated that the method could further enhance the system efficiency for collecting useful imaging information and therefore the image quality. In addition, this indirect optimum method could help the design of stationary SPECT system as well.

Conclusions and proposed future work are given in Chapter 7.

CHAPTER 2

IMAGING SYSTEM OPTIMIZATION

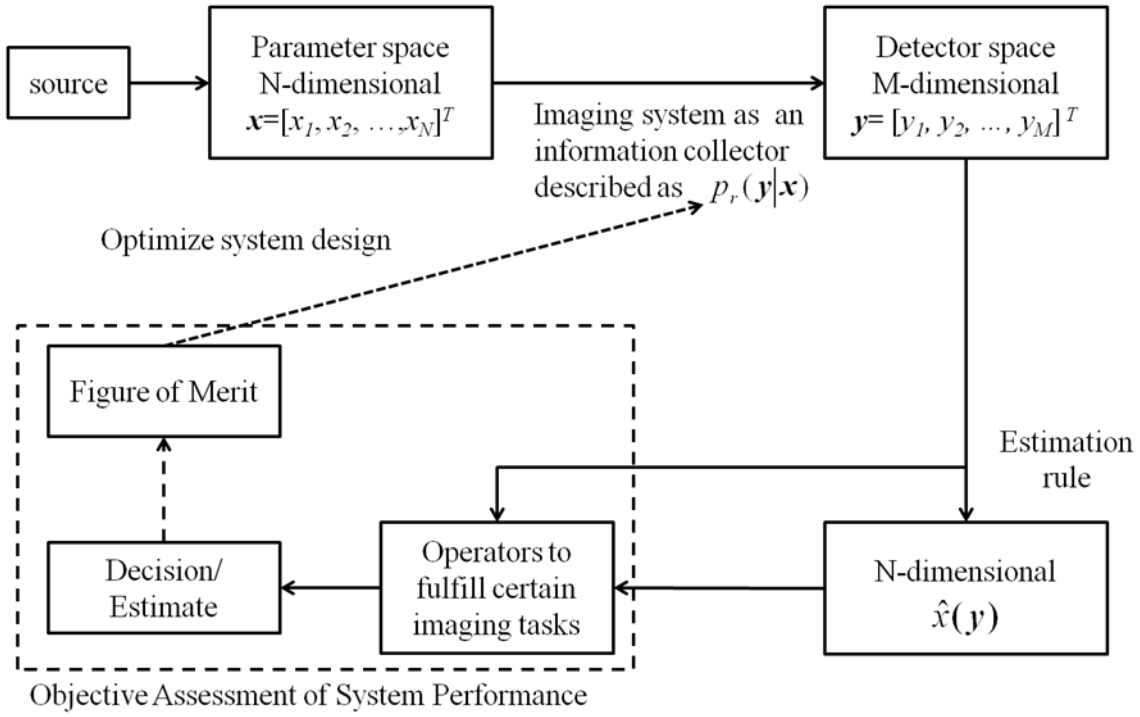

Figure 2.1 An illustration of imaging system optimization

In order to implement an optimization study for the performance of nuclear medical imaging systems, a set of requirements should be satisfied. First of all, the detection process and image reconstruction should be correctly modeled in mathematic forms. Subsequently, a specific evaluation method of the system performance should be defined. At last, a methodology should be implemented capable of determining the best performance against possible system and data acquisition parameters. To illustrate the generality of imaging system optimization, a block diagram is shown in Figure 2.1 as follows:

2.1 Object-Space Pixelation

While the true parameters in the object-space, e.g., the tracer intensities, are actually continuous functions, most researchers use a finite discrete distribution, denoted as $\boldsymbol{x}=[x_1, x_2, \ldots, x_N]^T$, to represent the object-space. In this case, the object-space is divided into N cubic voxels. The value of each element should in principle stand for the average of source intensity over the corresponding voxel. Although the aliasing error must be introduced in the pixelation,

we assume this approximation will not result in the information loss in the reconstructed images provided that source voxels are sufficiently closely spaced [3].

2.2 Detection Process

The unknown source could be imaged by an imaging system, which could be described as an information collector. The measurement data, denoted by vector $\boldsymbol{y}= [y_1, y_2, \ldots, y_M]^T$, represents the detected photons produced through radioactive decay. Note that the detector space is also discrete. The expectation of measurement follows a linear relationship with the true source in discrete form:

$$\bar{\boldsymbol{y}} = E[\boldsymbol{y}] = \boldsymbol{A} \cdot \boldsymbol{x} + \boldsymbol{r}, \tag{2.1}$$

where $E[\cdot]$ denotes the expectation operator. $\boldsymbol{A}$ is an $M \times N$ matrix that represents the discretized system response function (SRF) defined by the physical characteristics of the imaging system. In this matrix, element a_{mn} gives the probability of a gamma-ray emitted from the n'th source voxel and detected by the m'th detector pixel. $\boldsymbol{r}$ represents expectation of additive background events.

In emission tomography, the entire detection process is governed by a conditional probability $p_r(\boldsymbol{y}|\boldsymbol{x})$ mapping from unknown source space $\boldsymbol{x}$ to detector space $\boldsymbol{y}$. A measurement in one detector pixel could be described as a record of a photon emitted from the source as a product of radioactive decay. Then the detection in this detector pixel, considered as a collection of many measurements, could be modeled as Binomial distribution under the assumption that the probability of each measurement is a constant. Furthermore, suppose that this probability is quite small and independent, this model could be mathematically simplified as independent Poisson distribution [54]. Therefore, given the source distribution $\boldsymbol{x}$, the conditional probability for measurement data $\boldsymbol{y}$, also called the likelihood function, can be written as:

$$p_r(\boldsymbol{y}|\boldsymbol{x}) = \prod_m e^{-\bar{y}_m} \frac{\bar{y}_m^{\,y_m}}{y_m!}, \quad m = 1 \cdots M, \tag{2.2}$$

and

$$\bar{y}_m = \sum_n a_{mn} x_n + r_m. \tag{2.3}$$

Note that in SPECT problem, m in (2.2) specifies both a particular detector element and the sampling angle.

2.3 *Estimation Rule*

Nuclear imaging seeks to produce representations of the distribution of radioactive tracers within the body by detecting the photons emitted due to decay of the tracers. With measurements, the unknown source could be reconstructed following certain estimation rules. In principle, the reconstruction is equivalent to solve equation (2.1), taking into account the effects of the noise. Unfortunately, no straightforward methods exist for explicitly solving the exact imaging equations since system response function $\boldsymbol{A}$ is often singular and measurements are usually involved with random noise.

Maximum-likelihood (ML) estimate is a leading estimation method. For linear Poisson problem the measurement data $\boldsymbol{y}$ obeys independent Poisson distribution with mean equal to a linear transformation of the unknown parameter $\boldsymbol{x}$, the estimate of which could be reconstructed by maximizing the likelihood function (2.2), or equivalently, maximizing the natural logarithm of the likelihood function $L(\boldsymbol{x},\boldsymbol{y})$ [45, 46]:

$$\hat{\boldsymbol{x}}_{ML}(\boldsymbol{y}) = \underset{\mathbf{x}\geq 0}{\operatorname{argmax}}\left[p_r(\boldsymbol{y}|\boldsymbol{x})\right] = \underset{\mathbf{x}\geq 0}{\operatorname{argmax}}\left[L(\boldsymbol{x},\boldsymbol{y})\right] = \underset{\mathbf{x}\geq 0}{\operatorname{argmax}}\left[\sum_m y_m \log \bar{y}_m - \bar{y}_m\right]. \tag{2.4}$$

Ignoring constant terms independent with $\boldsymbol{x}$, the log-likelihood function $L(\boldsymbol{x},\boldsymbol{y})$ is defined as:

$$L(\boldsymbol{x},\boldsymbol{y}) = \log p_r(\boldsymbol{y}|\boldsymbol{x}) = \sum_m y_m \log \bar{y}_m - \bar{y}_m. \tag{2.5}$$

The advantage of ML estimate is that the corresponding estimator is always asymptotically unbiased and efficient, in the sense that it could achieve the lowest variance among all unbiased estimators. Moreover, if an efficient unbiased estimator really exists, it must be the ML estimator. In this case, the minimum mean square error could also be achieved [55].

However, in high resolution imaging applications, a common problem to ML estimate is ill-conditioned, i.e. the solutions are sensitive to small changes in the measurement data so that the estimates are of high variance. This problem could be observed through the effect that the variance of the image estimate may rapidly increase as the iterations proceed, especially in relatively low-counts detection situation [56]. The high variance may come from insufficient information acquired and thus the ML solutions are generally too noisy to be useful.

In practice, system resolution is often sacrificed to reduce noise in the images. It could be realized by using *a priori* information and smoothness constraints. The corresponding estimate is called post-filtered penalized maximum-likelihood estimate [57], given by:

$$\begin{cases} \hat{\boldsymbol{x}}_{PML}(\boldsymbol{y}) = \underset{\mathbf{x}\geq 0}{\arg\max}\left[L(\boldsymbol{x},\boldsymbol{y}) - \beta \cdot R(\boldsymbol{x})\right], \\ \text{and then} \\ \hat{\boldsymbol{x}}_{PF-PML} = \boldsymbol{F} \cdot \hat{\boldsymbol{x}}_{PML}(\boldsymbol{y}), \end{cases} \tag{2.6}$$

where $R(\boldsymbol{x})$ is a scalar function that selectively penalizes certain undesired features in the unknown image and β is a non-negative parameter that controls the degree of regularization. In our study, we usually use a quadratic roughness penalty function defined as [58]:

$$R(\boldsymbol{x}) = \sum_j \frac{1}{2} \sum_k w_{jk} \phi(x_j - x_k), \tag{2.7}$$

where w_{jk} s are the weighting factors that are non-zero for the pairs of immediate neighbors, and

$$\phi(\theta) = \theta^2 / 2. \tag{2.8}$$

$\boldsymbol{F}$ represents an $N \times N$ post-filtering operator. We typically use Gaussian filter matrix whose columns are defined as spatially shifted 3-D Gaussian functions centered at the related voxel positions with uniform full-width at half-maximum (FWHM). Then the post-filtering could be interpreted as a convolution operation. Both β and $\boldsymbol{F}$ can adjust the balance between the noise level and resolution achievable in reconstructed images.

2.4 Task-Based Objective Assessment of System Performance

In the system optimization, regardless of the imaging modality or the image-processing algorithm, a key question is how to define whether an imaging system performs better or not. It is generally agreed that the objective assessment of system performance must be task-based. The imaging tasks are generally divided into classification task and estimation task. Normally one could use some operators on either measurement or reconstructed images to fulfill certain imaging task: make a decision or estimate some parameters of interest. Based on them, one or more figures of merit are used to quantify how well the system performs the task [48]. This task-based objective assessment method provides rigorous definition of image quality, which could be used in system optimization.

2.4.1 Classification Task

In a classification task, there exist a finite number of classes. The goal is to decide to which class the image belongs. If only two classes exist, it is often referred to as a detection task. The simplest example is the detection of an abnormal feature for deciding whether there is a tumor or not.

A necessary component in classification is the specification of the observer to perform the task, which could be considered as an operator to make a decision using either measurement or reconstructed images. Some trained professionals, such as radiologists or cardiologists, are usually used as human observers in medical imaging, but the cost is so high that naive observers are also employed to aid in the assessment of image quality. In the study, an observer views a few images and then states whether images from one system are better or worse than another system. Of course human observer is irreproducible and subjective, and the time required is often prohibitively long. Instead, we could use some numerical observers which could be calculated unambiguously. As a common example, Hotelling observer is defined as a linear observer in detection task that maximizes the signal-to-noise ratio (SNR) of the observer's confidence ratings [59]. As seen in (2.9), this observer first computes a scalar product of the image vector $\hat{\boldsymbol{x}}$ with another vector $\boldsymbol{w}_{Hot}$ called the template or discriminant function:

$$t = \boldsymbol{w}_{Hot}^{T} \cdot \hat{\boldsymbol{x}}. \tag{2.9}$$

Compare the scalar *t* with a confidence threshold, one could produce a decision. To maximize the SNR of the confidence ratings, the Hotelling observer implements knowledge of the first- and second-order statistics of the image data that:

$$\boldsymbol{w}_{Hot} = \left[\mathrm{cov}\{\boldsymbol{y}\}\right]^{-1} \cdot \boldsymbol{A} \cdot \Delta E(\hat{\boldsymbol{x}}), \tag{2.10}$$

where $\mathrm{cov}\{\boldsymbol{y}\}$ is the covariance matrix of the measurement, which is approximately the same under both signal-present and signal-absent hypotheses for weak signals. $\Delta E(\hat{\boldsymbol{x}})$ denotes the difference in the means of $\hat{\boldsymbol{x}}$ under both hypotheses.

Having specified the observer, another component is the figure of merit to evaluate the performance of the observer to fulfill the tasks. One commonly-used performance measure is receiver operating characteristic (ROC) curve, in which the tradeoff between false-positive fraction (FPF) and true-positive fraction (TPF) is plotted using an increasingly permissive detection threshold. The TPF is the probability that an observer will decide a signal, e.g., tumor, is present when it is really present. The FPF is the probability that an observer will decide a signal is present when it is not actually present.

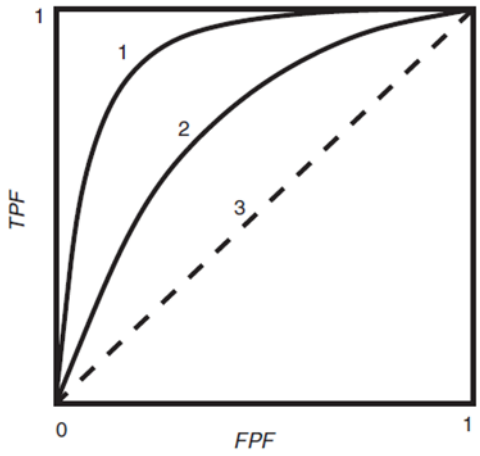

Figure 2.2 an example of ROC curves

In order to estimate the ROC curve for an observer, one could perform a two-alternative forced choice experiment in which the observer is presented two images and is forced to decide which image contains the signal. We have *a priori* knowledge that one of the images does not contain the signal and the other does. After presentation of a large number of pairs of images, the fraction of pairs where the observer correctly identified the abnormal image could be used to estimate the ROC curves with variable confidence thresholds. Commonly we use this figure-of-merit for assessing and comparing imaging system performances. Figure 2.2 shows an example of three ROC curves. In this figure, curve 1 owns a higher TPF at the same FPF compared with other two curves. Therefore, this curve is considered to exhibit the best performance among the three curves. Another generally applicable figure-of-merit is the area under the ROC curve (AUC) [1, 59]. A common surrogate performance index used in detection task, however, is a signal-to-noise ratio, which is defined as the ratio of the squared difference between the means of the two ratings under each class to the standard deviation of the ratings [47]. That is:

$$SNR \equiv \frac{2 \cdot \left[\langle t \rangle_1 - \langle t \rangle_0\right]^2}{\mathrm{var}_1(t) + \mathrm{var}_2(t)}, \tag{2.11}$$

where $\langle t \rangle_j$ is the expectation of the test statistic when hypothesis H_j is true (for example, H_0 is signal-absent and H_1 is signal-present), and $\mathrm{var}_j(t)$ is the corresponding variance. For Gaussian decision-variable data, the AUC is a simple monotonic function of the SNR.

2.4.2 Estimation Task

Estimation task involves the use of images to determine the value of one or more numerical parameters of interest. For example, the average activity level in a ROI could be quantified to

estimate the corresponding tumor uptake, growth rate and the efficacy of therapeutic interventions [60]. In our study, we focus on the estimation task and evaluate system performance based on certain figures-of-merit, such as tradeoff between spatial resolution and image variance for multiple voxels in the field-of-view (FOV) and/or tradeoff between mean bias and variance for ROI quantification, which reflect the balance between information per detected photon and sensitivity to detected photons.

Variance is a commonly-used statistical descriptor for measuring the precision, which is the spread of a set of estimate $\hat{\boldsymbol{x}}$ due to the statistical sampling [1, 61]. Suppose random estimate $\hat{\boldsymbol{x}}$ has the expected value $E(\hat{\boldsymbol{x}})$, the variance is defined as the expected value of the squared difference between the estimate's realization and the estimate's mean, given by:

$$\mathrm{var}(\hat{\boldsymbol{x}}) = E[(\hat{\boldsymbol{x}} - \boldsymbol{E}(\hat{\boldsymbol{x}}))^2]. \tag{2.12}$$

Brute-force Monte Carlo (MC) simulation method is a popular way to estimate the imaging noise characteristics by generating N realizations of noisy measurements which follow the identical and independent distribution. Then one can reconstruct the corresponding image estimates, denoted by $[\hat{\boldsymbol{x}}^1, \hat{\boldsymbol{x}}^2, \cdots, \hat{\boldsymbol{x}}^N]$, and the variance could be approximated as [54, 62]:

$$\mathrm{var}(\hat{\boldsymbol{x}}) \approx \frac{1}{N-1}\sum_{i=1}^{N}\left(\hat{\boldsymbol{x}}^{\boldsymbol{i}} - \frac{1}{N}\sum_{i=1}^{N}\hat{\boldsymbol{x}}^{\boldsymbol{i}}\right)^2, \tag{2.13}$$

where i denotes i'th noisy realization and N is the total number of realizations. For large size of sample N, this approximated variance approaches the true variance defined in (2.12). Therefore, a great number of realizations are usually necessary for sufficiently accuracy.

Difference from variance, bias measures the accuracy, which is the difference between the average estimate and the true parameter being estimated [1, 61], given by

$$Bias(\hat{\boldsymbol{x}}) = E(\hat{\boldsymbol{x}}) - \boldsymbol{x}. \tag{2.14}$$

An estimator with zero bias is called unbiased. Bias could also be estimated by the brute-force MC simulation using sample average instead of true average. The approximated bias is given by [54, 62]:

$$Bias(\hat{\boldsymbol{x}}) \approx \frac{1}{N}\sum_{i=1}^{N}\hat{\boldsymbol{x}}^{\boldsymbol{i}} - \boldsymbol{x}. \tag{2.15}$$

In practice, one could use a single value of mean percentage bias to evaluate the bias of noisy reconstructed images in the region of interest [63]. This value is calculated by comparing the reconstructed image averaged through all noisy realizations to the true image (source) in the ROI, given by:

$$Bias \approx \frac{1}{J}\sum_{j=1}^{J}\left\{\left(\frac{1}{N}\sum_{i=1}^{N}\hat{x}_j^i - x_j\right)\Big/ x_j\right\} \qquad j \in ROI. \tag{2.16}$$

where $\hat{x}_j^i$ is the value of the j'th voxel in the reconstruction of the i'th noisy realization, and x_j is the value of the j'th voxel in the true source. J is the total number of voxels in the ROI.

In addition to bias, spatial resolution is another basic measure of image accuracy for single voxel intensity estimation. It could be thought of as the ability of a medical imaging system to accurately depict two distinct events in the object-space. Therefore a high resolution medical imaging system is characterized by low smearing, whereas a low resolution system is characterized by high smearing. Given an estimator $\hat{\boldsymbol{x}}_j$, the corresponding local impulse response is commonly used to quantify the spatial resolution property, defined as [64]:

$$\boldsymbol{l}_j(\hat{\boldsymbol{x}}) = \frac{\partial E[\hat{\boldsymbol{x}}]}{\partial x_j} = \left[\frac{\partial E[\hat{\boldsymbol{x}}_1]}{\partial x_j}, \frac{\partial E[\hat{\boldsymbol{x}}_2]}{\partial x_j}, \cdots, \frac{\partial E[\hat{\boldsymbol{x}}_N]}{\partial x_j}\right]^T. \tag{2.17}$$

Then its full-width at half-maximum (FWHM) gives the minimum distance two point sources must be separated in space in order to appear as two points rather than one in the reconstructed images, seen in Figure 2.3 [4].

In order to estimate the local impulse response, one could still resort to the brute-force MC simulation by replacing the expectation in (2.17) by the sampling average. However, in emission tomography, several investigators have observed that the true mean of a likelihood-based estimator is approximately equal to the value of the estimator using noiseless data [56, 58, 62, 65]:

$$E[\hat{\boldsymbol{x}}(\boldsymbol{y})] \approx \hat{\boldsymbol{x}}[\bar{\boldsymbol{y}}(\boldsymbol{x})]. \tag{2.18}$$

where $\bar{\boldsymbol{y}}(\boldsymbol{x})$ is the expectation of the measurement $\boldsymbol{y}$, and is the function of the true source. This approximation is equivalent that the estimator is locally linear [64]. Substituting (2.18) into (2.17) yields the following definition of the linearized local impulse response:

$$\boldsymbol{l}_j(\hat{\boldsymbol{x}}) = \lim_{\delta \to 0} \frac{\hat{\boldsymbol{x}}[\bar{\boldsymbol{y}}(\boldsymbol{x} + \delta \cdot \boldsymbol{e}_j)] - \hat{\boldsymbol{x}}[\bar{\boldsymbol{y}}(\boldsymbol{x})]}{\delta}. \tag{2.19}$$

For brevity, "linearized" is usually omitted in the text later. In practice, the local impulse response could be approximated as:

$$\boldsymbol{l}_j(\hat{\boldsymbol{x}}) \approx \frac{\hat{\boldsymbol{x}}[\bar{\boldsymbol{y}}(\boldsymbol{x} + \delta \cdot \boldsymbol{e}_j)] - \hat{\boldsymbol{x}}[\bar{\boldsymbol{y}}(\boldsymbol{x})]}{\delta}. \tag{2.20}$$

By taking δ sufficiently small, one could obtain very accurate local impulse response. (2.20) leads to a much simpler calculation and only two noiseless reconstructions are required.

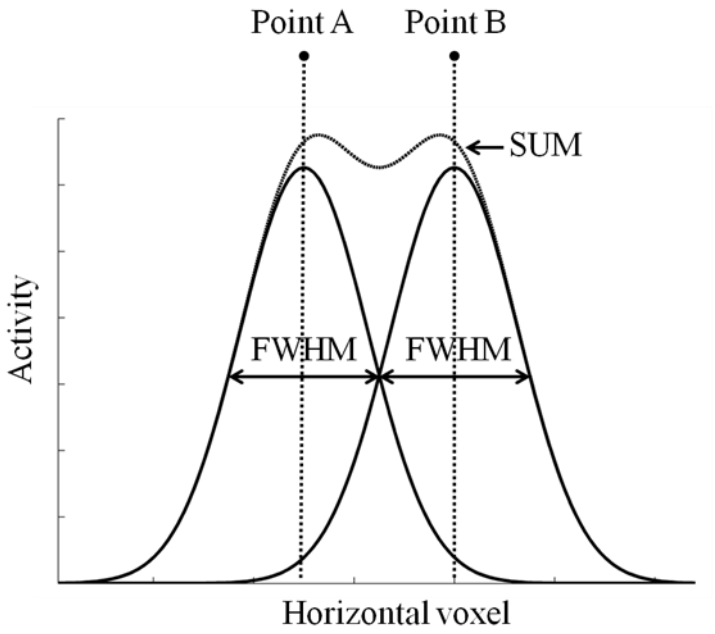

Figure 2.3 The limited resolution blurs activity from two discrete point sources (A and B), resulting in the apparent activity diffusion into regions without radioactivity. If these two point sources are separated by a distance of the FWHM, the summed activity has a modest decrease at the midpoint, suggesting that the original source consist of two points rather than one.

One important application of emission tomography is to quantify the tracer uptake in the region-of-interest (ROI). It is essential to use bias-variance tradeoffs to quantify the image quality inside the ROI [60, 66]. For single voxel estimation, however, this tradeoff has its drawbacks. One the one hand, one can always find an estimator with zero variance at a given voxel. On the other hand, the use of bias for a single voxel can be misleading that one could remove a constant bias even it is very large [67, 68]. In this case, the change of the activity value is more meaningful and thus we could use the resolution-variance tradeoff instead to quantify system performance.

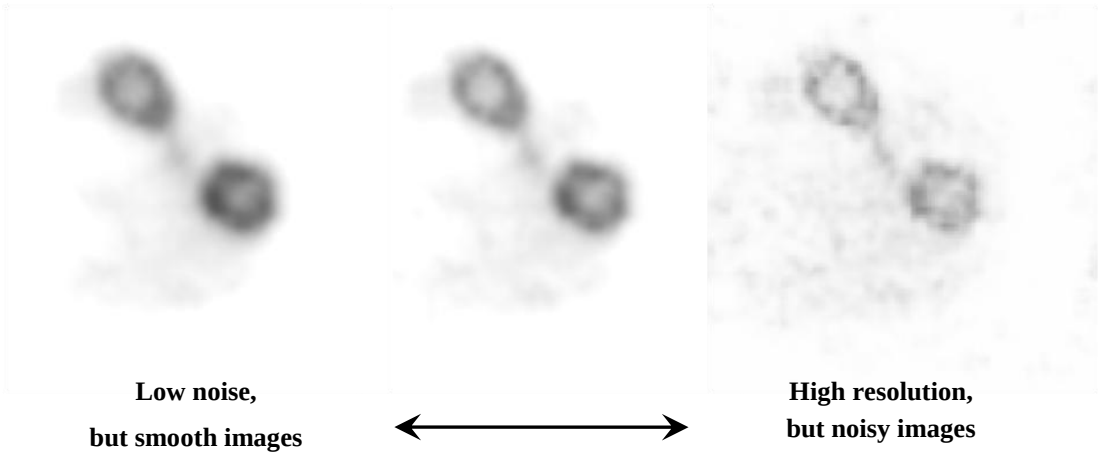

Figure 2.4 cross-sectional slice of mouse kidneys showing tradeoffs between reconstructed image noise and resolution

Clearly a good estimate should have simultaneously small bias and variance or small resolution and variance. Unfortunately, it is typically true that bias and variance of an estimate cannot be reduced at the same time and neither can resolution and variance. Rather, one can only reduce the bias or resolution at the expense of increasing the variance, and vice versa. Figure 2.4 illustrates the resolution and noise tradeoffs for the reconstructed images of mouse kidneys, which was imaged by the dual-head SPEM system [40]. Therefore, when evaluating the performance of an imaging system, one needs to consider both bias (resolution) and variance at the same time. Compared with a single value, these tradeoffs could provide a more complete performance assessment.

2.5 System Performance Optimization

In system optimization, it is desired to obtain the best performance for an imaging task, measured by a given figure-of-merit Ω:

$$\boldsymbol{q}_{optimum} == \underset{\boldsymbol{q}=[q_1,q_1,\cdots q_S]}{\operatorname{argmax}} [\Omega(\boldsymbol{q})], \tag{2.21}$$

where vector $\boldsymbol{q}=[q_1, q_2, \ldots, q_S]$ denotes the system parameters to be optimized. These parameters could include detector and aperture positions, detector intrinsic resolution, configurations of pinholes on each aperture and many more in SPECT imaging and ring diameter, axial size and thickness of the detector and more in PET imaging. The figure-of-merit Ω is chosen to represent how well the given imaging system to fulfill an imaging task or a class of tasks, which is typically a scalar or a vector function of $\boldsymbol{q}$. For example, one could choose to optimize an imaging system

for estimation tasks, based on performance indices such as spatial resolution [64, 69], image variance [58, 70] and their tradeoffs [71, 74], and/or the accuracy for ROI quantification [60, 66]. Alternatively, one could optimize a system for detection tasks based on the receiver-operation characteristics (ROC) of computer or human observers [75-83]. During the optimization process, system configurations could be adjusted, maybe iteratively, until the best figures-of-merit are obtained. In that case, the optimized system performance is considered to be achieved.

For high resolution imaging system, it is required that the practical optimization should be realized in real-time but the computational requirements may not be compatible with complexity of acquisition parameters, even with state-of-the-art performance parallel computing platform. On the one hand, algorithms of image reconstruction and system performance evaluation are often time-consuming. On the other hand, for typical imaging systems, the relationships between the figures-of-merit and system parameters are highly complex and often mathematically intractable. A generally accepted analytical approach for direct system optimization against the design parameters may not be available. Furthermore, a system optimization procedure typically requires a search through a vast (often infinite) number of possible system and sampling configurations. Thus it is almost impossible to use brute-force numerical approaches to search through the vast parameter space.

To address these questions, we have developed four approaches. Both vector modified uniform Cramer-Rao bound and non-uniform object-space pixelation approach have been developed for rapidly evaluating system performance indices, such as resolution-variance tradeoffs and bias-variance tradeoffs. Based on them, both adaptive angular sampling approach and indirect system optimization approach provide an efficient method for optimizing system configuration and/or sampling strategy within infinite parameter space. The series of analytical approaches developed through this thesis work could be used together to provide an efficient computation scheme to facilitate real-time system optimization.

In general, the optimized configuration should be dependent with the imaging task being performed and specific target being imaged. In practice an imaging system is usually employed for a variety of different imaging applications, such as diagnose of tumor, performance evaluation for different ROI quantifications and so on. Therefore, the practical optimization approach developed could be used along with variable hardware detection system to implement adaptive imaging, which could help to maximize the efficiency of collecting useful information from the unknown object.

CHAPTER 3
VECTOR MODIFIED UNIFORM CRAMER-RAO BOUND FOR IMAGING SYSTEM DESIGN

One of the key challenges for ultra-high resolution SPECT design is to achieve a reasonable tradeoff between resolution and detection efficiency, which reflects the balance between information per detected photon and total number of detected photons. As the most widely used approach for system performance estimation, Monte Carlo (MC) simulation allows detailed modeling and comprehensive evaluations for a given design. However, MC studies are often time-consuming, especially for evaluating the noise characteristics offered by an ultra-high resolution imaging system. Let's take the dual-head SPEM system developed in our lab as an example [40]. This system could offer an ultra-high spatial resolution capacity for a targeted local region inside mouse brain, which requires very small voxel sizes in the object-space, say 300 μm or below. If such voxel sizes are uniformly used throughout the entire object-space, the reconstruction would involve a tremendous amount of computation. Even with appropriate exploitation of the sparseness associated with multiple pinhole geometry, a single reconstruction typically takes almost one day to be completed using a 32-CPU cluster that we assembled for this application. As previously introduced, a large number of reconstructions are required in brute-force MC simulation for estimating image variance and bias, while two reconstructions are necessary for accurately estimating spatial resolution properties for single voxel. Therefore, the computation load involved in MC simulation is too tremendous to be practical in system performance evaluation. To overcome this problem, many analytical methods have been developed. Rentmeester et al. have presented an analytical model for optimizing an ultra-high resolution small animal SPECT system [69]. Bal et al. have reported a geometric-criterion based optimization method for optimizing the arrangement of pinholes [84]. Accorsi et al. have used signal-to-noise ratio (SNR) as a design criterion for selecting coded-aperture patterns [85]. In general, analytical methods often rely on simplified system models that lead to a greatly improved computation efficiency. This allows one to search through the multivariate system parameter space with a reasonable computation load.

SPECT systems may also be evaluated based on the fundamental tradeoffs between spatial resolution and image noise that can be achieved in reconstructed images. Such tradeoffs may be

derived analytically using the Cramer-Rao (CR) type bounds [86]. The ordinary CR bound is applicable to unbiased estimators only. This limits its use in imaging problems, since estimator bias is typically inevitable. The biased CR bound could apply to the estimator with particular bias gradient, but the derived estimator-dependent variance bound would confuse the effect of the imaging system to the system performance. The uniform Cramer-Rao bound (UCRB), developed by Hero et al. [67], provides the smallest attainable variance with any estimator whose bias gradient satisfies a certain constraint. Since bias gradient is an indirect measure of imaging resolution, one may plot the UCRB against the norm of the corresponding bias gradient vector. This leads to the so-called resolution-variance tradeoff curve. An example of using the UCRB for evaluating SPECT systems was presented by Hua et al. [87]. In this work, Compton cameras were compared with mechanically collimated gamma cameras for SPECT imaging applications. Furthermore, UCRB approach provides the optimum tradeoffs between voxel-wise resolution and variance, which is independent of the estimator used. One of the key problems of this approach is that the bias gradient norm constraint was inadequate for determining spatial resolution function. For example, point-spread functions with very different FWHM and FWTM values may satisfy the same bias gradient norm constraint [88]. Therefore, the relationship between bias gradient norm and variance may not represent the true tradeoff between imaging resolution and variance.

To overcome this limitation, Meng et al. have previously proposed a modified uniform Cramer-Rao bound (MUCRB) [72]. It follows the same notation as in Hero's formulation, but with a key difference. The MUCRB uses a resolution constraint that is applied on the mean gradient rather than the bias gradient of an estimator. The MUCRB offers two benefits over the original UCRB by Hero. Firstly, mean gradient is closely related to the linearized local impulse response (LIR) function [64]. The latter is widely used to describe spatial resolution property in images. Therefore, the MUCRB allows one to compare the true resolution-variance tradeoff that is attainable with given imaging system designs. Secondly, for linear Gaussian and linear Poisson problems, the MUCRB can be achieved (asymptotically) with post-filtered penalized maximum likelihood estimators (PF-PML) that are widely used in routine practice. In our previous studies, the use of the MUCRB has provided reasonable predictions in comparing different SPECT aperture designs [31, 72].

However, the formulation of the MUCRB is limited to the scalar estimation problem. It gives the minimum variance attainable with any estimator of a scalar function of a vector parameter, e.g. the tracer uptake in a single voxel in an image. From a practical view point, one may be more interested in the estimation of a vector function of underlying unknown parameters. For example, SPECT systems are normally designed to provide simultaneous estimate of tracer uptakes in

multiple voxels. In this section, we extended the MUCRB into the vector estimation case. This vector MUCRB allows the evaluation of the minimum achievable total (or average) variance over an arbitrary set of voxels, given that the spatial resolution functions achieved at these voxels satisfy a certain constraint. This development provides a general form of UCRB that incorporates the original UCRB by Hero et al. and the scalar MUCRB as special cases. This offers a more meaningful way to compare the resolution-variance tradeoffs achieved with different systems.

In the following text, we will provide a brief review of the development of Cramer-Rao type bound, followed by the derivation of the vector MUCRB. The use of the vector MUCRB for SPECT system design will be demonstrated with a Monte Carlo example.

3.1 Development of Cramer-Rao Type Bound

3.1.1 Cramer-Rao Bound

In estimation theory and statistics, Harald Cramér and Calyampudi Radhakrishna Rao derived a lower bound on the variance of unbiased estimators, which is named as the Cramer-Rao bound [86]. Suppose $\boldsymbol{x}$ is a vector of unknown deterministic parameters to be estimated. Observed random measurement $\boldsymbol{y}$ obeys the likelihood function $p_r(\boldsymbol{y}|\boldsymbol{x})$. Given the statistical model, the Cramer-Rao bound states that the covariance matrix of unbiased estimator $\hat{\boldsymbol{x}}$ of $\boldsymbol{x}$ satisfies:

$$Cov(\hat{\boldsymbol{x}}) \geq \boldsymbol{J}^{-1} \equiv \left\{ -E\left[\nabla_x^2 \ln p_r(\boldsymbol{y}|\boldsymbol{x})\right]\right\}^{-1} \equiv \left\{ -E\left[\nabla_x^2 L(\boldsymbol{x},\boldsymbol{y})\right]\right\}^{-1}, \tag{3.1}$$

where matrix inequality is understood as that the matrix $Cov(\hat{\boldsymbol{x}}) - \boldsymbol{J}^{-1}$ is positive semi-definite, and $E[\cdot]$ denotes the expectation operator. $\boldsymbol{J}$ is the Fisher information matrix (FIM), and L is the natural logarithm of the likelihood function $p_r(\boldsymbol{y}|\boldsymbol{x})$ [58].

The bound relies on a weak regularity condition that $\nabla_x \ln p_r(\boldsymbol{y}|\boldsymbol{x})$ exists and is finite, which could be satisfied in linear Poisson problem. With this noise model, the log-likelihood function L is defined in (2.5) and the corresponding element of Fisher information matrix could be explicitly defined as:

$$\boldsymbol{J}_{ij} = -E\left\{\left[\frac{\partial^2}{\partial x_i \partial x_j} L(\boldsymbol{x},\boldsymbol{y})\right]\right\} = \sum_m \frac{a_{mi} a_{mj}}{\bar{y}_m}, \tag{3.2}$$

where a element a_{mn} in the system response function $\boldsymbol{A}$ gives the probability of a gamma-ray emitted from the n'th source voxel and detected by the m'th detector pixel. Therefore equation (3.2) indicates that each column of FIM measures the correlation between the responses of the

system to gamma ray emission at the related voxel and at all voxels in object-space, weighted by the mean value of the projection data. This equation can be re-written in the matrix form as:

$$\boldsymbol{J} = \boldsymbol{A}^T \cdot diag\left\{\frac{1}{\bar{y}_m}\right\} \cdot \boldsymbol{A}. \tag{3.3}$$

For single voxel estimation, the minimum variance of an unbiased estimator $\hat{x}_j$ for *j*'th voxel intensity is given by

$$\text{var}(\hat{x}_j) \geq \boldsymbol{e}_j^T \cdot \boldsymbol{J}^{-1} \cdot \boldsymbol{e}_j, \tag{3.4}$$

where vector $\boldsymbol{e}_j$ is the unit vector related to the *j*'th voxel.

Since Fisher information matrix is implicitly determined only by the physical characteristics of the imaging system and true source function, this minimum variance bound is estimator-independent and thus can be used to find the fundamental limitations of an imaging system. Note that variance bound is only applied to unbiased estimators and could be asymptotically achieved by the maximum-likelihood estimator. As previously discussed, ML estimator is usually too noisy to be useful, which indicates that the Cramer-Rao bound is not practical in our application.

3.1.2 Biased Cramer-Rao Bound

In high resolution imaging application, it is often necessary to add some bias to estimator in order to reduce the noise. Then the biased Cramer-Rao bound derives the bound on the covariance matrix of biased estimators $\hat{\boldsymbol{x}}$ of the parametric vector $\boldsymbol{x}$:

$$Cov(\hat{\boldsymbol{x}}) \geq (\boldsymbol{I} + \nabla \boldsymbol{b}) \cdot \boldsymbol{J}^{-1} \cdot (\boldsymbol{I} + \nabla \boldsymbol{b}), \tag{3.5}$$

with a given bias vector $\boldsymbol{b} = E(\hat{\boldsymbol{x}}) - \boldsymbol{x}$. $\boldsymbol{I}$ is the identity matrix. Note that if the estimator is unbiased and thus $\nabla \boldsymbol{b} = 0$, the equation (3.5) reduces to classical Cramer-Rao bound (3.1). So the unbiased version of the Cramer-Rao bound is a special case of the biased Cramer-Rao bound.

For single voxel estimation, suppose a biased estimator $\hat{x}_j$ for *j*'th voxel intensity has mean value of $m_j = E(\hat{x}_j)$, the corresponding biased Cramer-Rao bound has the following form:

$$\begin{aligned} \text{var}(\hat{x}_j) &\geq (\nabla b_j + \boldsymbol{e}_j)^T \cdot \boldsymbol{J}^{-1} \cdot (\nabla b_j + \boldsymbol{e}_j) \\ &= \nabla m_j{}^T \cdot \boldsymbol{J}^{-1} \cdot \nabla m_j, \end{aligned} \tag{3.6}$$

where bias $b_j = m_j - x_j$. Thus the gradient of the estimator bias function ∇b_j (equivalently bias gradient vector) and the gradient of the estimator mean function ∇m_j (equivalently mean gradient vector) are related by:

$$\nabla m_j = \nabla(b_j + x_j) = (\nabla b_j + \boldsymbol{e}_j). \tag{3.7}$$

Note that this bound can be less than the unbiased Cramer-Rao bound (3.4), at the price of the existence of bias. The item of estimator mean function in (3.6) indicates that the variance bound depends not only on the imaging system itself but also on the estimator to create the image, which could confuse the effect of the imaging system to the system performance. Furthermore, given a particular bias gradient function, (3.6) could compute the corresponding minimum variance. However, it is usually impractical to calculate a desired bias gradient and therefore the biased Cramer-Rao bound cannot bound variance of estimators simultaneously. In some cases, it may be difficult to determine the optimum estimator among those with different but comparable bias gradient vectors and variance bounds.

3.1.3 Uniform Cramer-Rao Bound

In order to quantify the fundamental performance of an imaging system for parametric estimation problem, Hero et al. developed the uniform Cramer-Rao bound [67], which provides an effective way to compare different imaging systems. For single voxel estimation, let $\hat{x}_j$ be an estimator for *j*'th voxel intensity in the object space. The idea behind the uniform Cramer-Rao bound is that for any estimator whose bias gradient norm is equal to or less than a fixed positive value δ

$$\left\| \nabla_x b(\hat{x}_j) \right\|_{\boldsymbol{C}} = \left\| \nabla_x E(\hat{x}_j) - \boldsymbol{e}_j \right\|_{\boldsymbol{C}} \le \delta, \tag{3.8}$$

the minimum achievable variance is bounded as

$$\mathrm{var}(\hat{x}_j) \ge \lambda^2 \cdot \boldsymbol{e}_j^{\boldsymbol{T}} \cdot [C^{-1} + \lambda]^{-1} \cdot \boldsymbol{J} \cdot [C^{-1} + \lambda]^{-1} \cdot \boldsymbol{e}_j, \tag{3.9}$$

where λ is determined by the unique positive solution of the following equation:

$$\boldsymbol{e}_j^T \cdot \left[\boldsymbol{C}^{-1} + \lambda \boldsymbol{J}\right]^{-1} \cdot \boldsymbol{C}^{-1} \cdot \left[\boldsymbol{C}^{-1} + \lambda \boldsymbol{J}\right]^{-1} \cdot \boldsymbol{e}_j = \delta^2 \tag{3.10}$$

The bias gradient of the efficient estimator that achieves this bound is given by:

$$\left[\nabla_x b(\hat{x}_j)\right]_{\min} = -\boldsymbol{C}^{-1} \cdot \left[\boldsymbol{C}^{-1} + \lambda \boldsymbol{J}\right]^{-1} \cdot \boldsymbol{e}_j. \tag{3.11}$$

where $\boldsymbol{C}$ is a symmetric and positive definite weighting matrix defined by the user. In actual implementation, the efficient estimator function could not be actually solved in (3.11). Instead, one could vary the threshold δ and compute the corresponding minimum achievable variance using (3.10) and (3.9) to derive the tradeoffs between variance bound and bias-gradient norm. All the other estimators satisfying the bias-gradient norm constraint (3.8) must not have lower variance. If an estimator lies on the curve, variance can be further reduced only at the price of

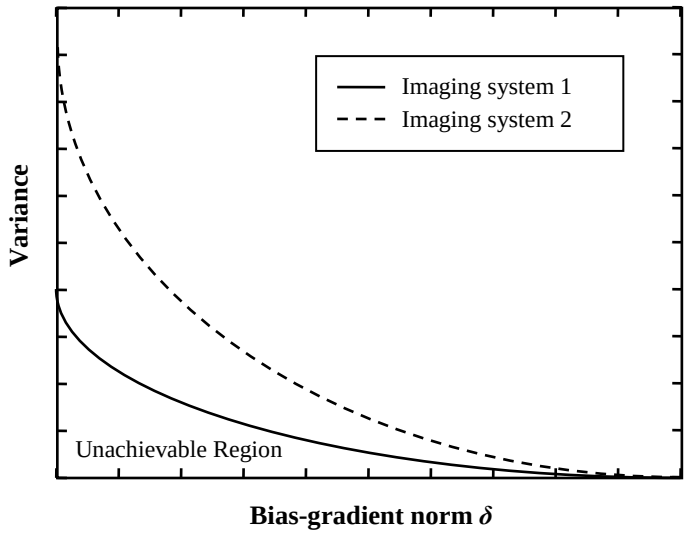

Figure 3.1 comparison of tradeoffs between bias-gradient norm and variance derived for two imaging systems

increased estimator bias and vice versa [67]. Furthermore, the variance bound is only a function of the Fisher information matrix $\boldsymbol{J}$ and the given bias-gradient norm δ, and independent with any particular choice of estimation algorithm. Therefore, the derived tradeoff curve assesses fundamental performance limit of the imaging system. In practice, it is important to point out that the tradeoff curves can be generated by continuously varying intermediate variable λ over the range $(0, \infty)$, instead of solving the nonlinear equation (3.10). Figure 3.1 gives a simple example of comparing two tradeoffs between bias-gradient norm and variance for different imaging systems. In this figure, imaging system 1 could obtain lower variance bound at any given bias-gradient norm and thus own the better performance than imaging system 2. In addition, solid line gives the "ideally" best tradeoffs between bias-gradient norm and variance, and the region under the curve could not be achieved by any other estimators for the imaging system 1. Therefore, the curve quantifies the fundamental system performance limit which is independent with the choice of estimators.

In (3.8) the mean gradient vector $\nabla E(\hat{x}_j) = \left[\frac{\partial E[\hat{\boldsymbol{x}}_j]}{\partial x_1}, \frac{\partial E[\hat{\boldsymbol{x}}_j]}{\partial x_2}, \cdots, \frac{\partial E[\hat{\boldsymbol{x}}_j]}{\partial x_N} \right]^T$ measures the sensitivity into a single mean reconstructed voxel due to variation in all source voxels. In comparison, the local impulse response $\frac{\partial E[\hat{\boldsymbol{x}}]}{\partial x_j} = \left[\frac{\partial E[\hat{\boldsymbol{x}}_1]}{\partial x_j}, \frac{\partial E[\hat{\boldsymbol{x}}_2]}{\partial x_j}, \cdots, \frac{\partial E[\hat{\boldsymbol{x}}_N]}{\partial x_j} \right]^T$ is defined as the sensitivity to all the mean reconstructed voxels due to a perturbation in a single source voxel. It is easy to find that the mean gradient is closely related to the local impulse response in

definition. In our study we assume that mean gradient vector is equivalent with local impulse response, which was validated in [73]. Therefore mean gradient vector can also be used to quantify the resolution property as local impulse response.

As a result, bias-gradient vector norm constraint (3.8), which measures the difference between actual and ideal mean gradient vector, in some sense quantifies the degree of blurring in reconstructed images. However, this norm constraint is inadequate to confine the system resolution and thus not the right resolution measure. In practice, most mean gradient vector is significantly different from the ideal point spread function. So there may exist many estimators which have the same bias gradient norm, but the resolution property is obviously different [88, 89]. A simple example is shown in Figure 3.2. Therefore the relationship between bias gradient norm and variance may not represent the true resolution-variance tradeoff.

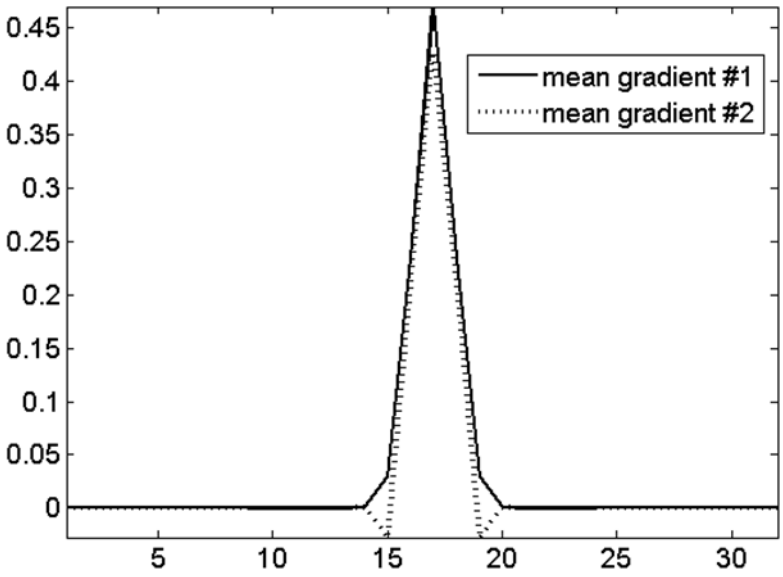

Figure 3.2 cross-sectional slices through two mean gradient vectors as a function of pixel location. Their associated bias gradients both have the same norm, but their resolution properties are obviously different

3.1.4 Modified Uniform Cramer-Rao Bound (MUCRB)

To overcome this limitation, Meng et al. [72] has proposed a modified uniform Cramer-Rao bound (MUCRB). In this approach, a new resolution constraint is applied to measure the difference between mean gradient vector $\nabla E(\hat{x}_j)$, expressed by $\boldsymbol{g}_j$, and desired local impulse response $\boldsymbol{f}_j$ instead of the ideal local impulse response:

$$\left\| \boldsymbol{g}_j - \boldsymbol{f}_j \right\|_{\boldsymbol{C}} \le \delta. \tag{3.12}$$

This new resolution constraint imposes a stronger limitation on the actual shape of the mean gradient vector. If the threshold is small, the resulting mean gradient vector can only deviate slightly from the target $\boldsymbol{f}_j$. All estimators that satisfy constraint (3.12) with a small threshold δ should produce very similar spatial resolution properties, regardless of the physical system configuration and the estimation method used. So this allows a more meaningful constraint on the achievable spatial resolution property.

Minimum variance for any biased estimator is given by biased Cramer-Rao bound (3.6). With (3.12) in place, the optimum mean gradient vector that gives the constrained minimum achievable variance could be described as

$$\boldsymbol{g}_{optimum} = \underset{\left\| \boldsymbol{g}_j - \boldsymbol{f}_j \right\|_{\boldsymbol{C}} \le \delta}{\arg\min} \left\{ \boldsymbol{g}_j^T \cdot \boldsymbol{J}^{-1} \cdot \boldsymbol{g}_j \right\}. \tag{3.13}$$

In mathematical optimization, Karush–Kuhn–Tucker (KKT) conditions [90] provides a strategy for finding the local minima of the biased Cramer-Rao bound subject to the resolution constraints (3.12). KKT multiplier λ is introduced to form:

$$f_{KKT} = \boldsymbol{g}_j^T \cdot \boldsymbol{J}^{-1} \cdot \boldsymbol{g}_j + \lambda \left\{ \left(\boldsymbol{g}_j - \boldsymbol{f}_j \right)^T \cdot \boldsymbol{C} \cdot \left(\boldsymbol{g}_j - \boldsymbol{f}_j \right) - \delta^2 \right\}, \quad \lambda \ge 0 \tag{3.14}$$

Since f_{kkt} is strictly concave, it has a unique minimum. Let the gradient of f_{KKT} with respect to $\boldsymbol{g}_j$ equal to zero and we have

$$\nabla_{\boldsymbol{g}_j} f_{KKT} = \boldsymbol{J}^{-1} \cdot \boldsymbol{g}_j + \lambda \cdot \boldsymbol{C} \cdot \left(\boldsymbol{g}_j - \boldsymbol{f}_j \right) = 0. \tag{3.15}$$

Then the minimum of f_{KKT} is achieved with the optimum mean gradient, given by

$$\boldsymbol{g}_{optimum} = \boldsymbol{J} \cdot \left[\boldsymbol{J} + \left(\lambda \cdot \boldsymbol{C} \right)^{-1} \right]^{-1} \cdot \boldsymbol{f}_j. \tag{3.16}$$

Substituting $\boldsymbol{g}_{optimum}$ into (3.13), the associated constrained minimum attainable image variance at the *j*'th voxel could be described as:

$$Var(\hat{\boldsymbol{x}}_j) \ge \boldsymbol{f}_j^T \cdot \left[\boldsymbol{J} + \left(\lambda \cdot \boldsymbol{C} \right)^{-1} \right]^{-1} \cdot \boldsymbol{J} \cdot \left[\boldsymbol{J} + \left(\lambda \cdot \boldsymbol{C} \right)^{-1} \right]^{-1} \cdot \boldsymbol{f}_j. \tag{3.17}$$

It is easy to find that this bound is also estimator-independent. Furthermore, the optimum mean gradient in (3.16) and corresponding minimum achievable variance in (3.17) derive the fundamental performance tradeoffs over other estimators.

In addition, the small threshold δ in resolution constraint (3.12) indicates that $\boldsymbol{g}_{optimum}$ should be very similar as the desired resolution $\boldsymbol{f}_j$. Therefore, it could be demonstrated in (3.16) that the

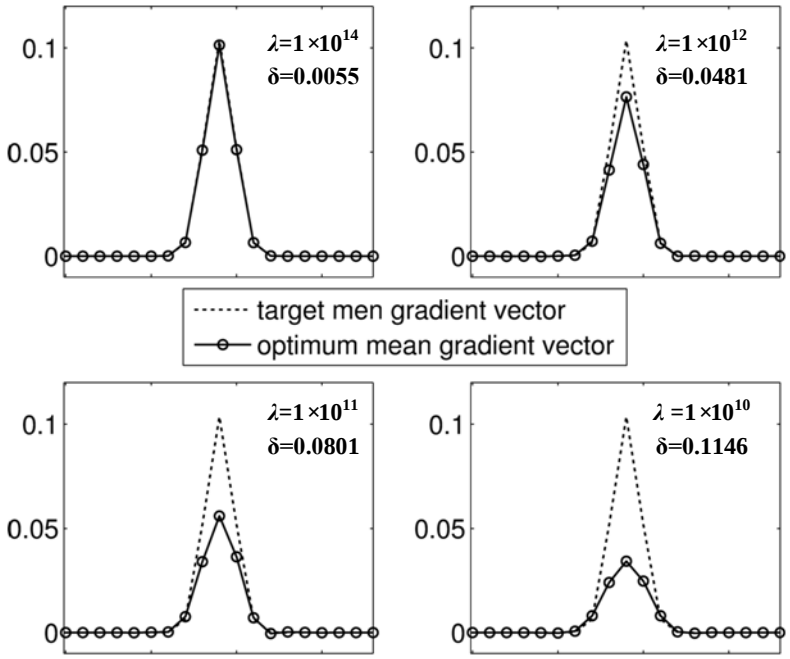

Figure 3.3 Optimum (circles and solid curves) and target (dashed curves) mean gradient vectors for different λ. The resulting resolution constraint norms become smaller when larger λ is used.

value of λ should be very large and resolution property is mainly determined by $\boldsymbol{f}_j$. Figure 3.3 gives a simple example to illustrate the change of optimum mean gradient vectors of efficient estimator as a function of λ. The FWHM of target mean gradient vector (local impulse response) is fixed to be 2 voxels and weighting matrix $\boldsymbol{C}$ is chosen as identity matrix. Clearly the shape of optimum mean gradient vector approaches the target one when increasing λ and for the large value of λ (1×10^{14}) resolution constrains could be satisfied with very small norm δ of 0.0055.

Varying the full-width-at-half-maximum (FWHM) of the desired filter function $\boldsymbol{f}_j$ instead of λ, we could obtain the fundamental tradeoff curve of resolution and achievable minimum variance, by which we could compare performances of different system models. In the previous studies, the MUCRB approach has provided reasonable predictions in comparing different SPECT aperture designs [72].

3.2 Derivation of Vector Modified Uniform Cramer-Rao Bound

The vector modified UCRB (3.17) is limited into the scalar case, and one can derive the optimum voxel-wise resolution-variance tradeoffs attainable with any estimator of a scalar function of a vector parameter, e.g. the tracer uptake in a single voxel in an image. From a practical view point, however, we are normally interested in estimating multiple components (or a

vector function) simultaneously for evaluating an imaging system. For example, an ultra-high resolution SPECT may be designed to provide an optimized performance for imaging multiple voxels in a region-of-interest (ROI). Therefore, it is desirable to extend the original MUCRB into vector estimation case when the imaging task is to estimate a vector function of the underlying parameters. This vector MUCRB allows the evaluation of average or collective resolution-variance tradeoffs over an arbitrary set of voxels at a given resolution and offers a more meaningful way to compare the performances of different systems.

For the estimation problem outlined in the previous section, the mean gradient matrix of a vector estimator $\hat{\boldsymbol{x}}$ is defined as [73]:

$$\boldsymbol{G} = \frac{\partial E[\hat{\boldsymbol{x}}]}{\partial \boldsymbol{x}} = \begin{bmatrix} \frac{\partial}{\partial \boldsymbol{x}_1} E[\hat{\boldsymbol{x}}_1] & \frac{\partial}{\partial \boldsymbol{x}_2} E[\hat{\boldsymbol{x}}_1] & \cdots & \frac{\partial}{\partial \boldsymbol{x}_N} E[\hat{\boldsymbol{x}}_1] \\ \frac{\partial}{\partial \boldsymbol{x}_1} E[\hat{\boldsymbol{x}}_2] & \frac{\partial}{\partial \boldsymbol{x}_2} E[\hat{\boldsymbol{x}}_2] & \cdots & \frac{\partial}{\partial \boldsymbol{x}_N} E[\hat{\boldsymbol{x}}_2] \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial}{\partial \boldsymbol{x}_1} E[\hat{\boldsymbol{x}}_N] & \frac{\partial}{\partial \boldsymbol{x}_2} E[\hat{\boldsymbol{x}}_N] & \cdots & \frac{\partial}{\partial \boldsymbol{x}_N} E[\hat{\boldsymbol{x}}_N] \end{bmatrix}. \tag{3.18}$$

Each row of $\boldsymbol{G}$ is the mean gradient vector for a given voxel in the image and each column is the local impulse response function as defined in (2.17). If the mean gradient matrix $\boldsymbol{G}$ is symmetric, the mean-gradient vector is equal to the local impulse response for the same voxel. So the mean gradient matrix could be used as a comprehensive description of resolution property in the images. Based on it, an "average" resolution constraint for a subset of estimators is defined as

$$s = tr\left\{\boldsymbol{W} \cdot (\boldsymbol{G} - \boldsymbol{F}) \cdot \boldsymbol{C} \cdot (\boldsymbol{G} - \boldsymbol{F})^T \cdot \boldsymbol{W}^T\right\} \le \gamma, \tag{3.19}$$

where $\boldsymbol{F}$ is the desired spatial resolution function in reconstructed images, each row of which gives the desired (or target) mean-gradient vector for a given voxel. $\boldsymbol{W}$ is a nonnegative definite weighting matrix. The scalar s basically measures a weighted average over all Euclidean distances between true and target mean gradient vectors for all voxels in a user-defined region. For a task of optimizing SPECT systems for imaging a ROI, $\boldsymbol{W}$ is usually defined as:

$$\boldsymbol{W} = diag\{w_j\}, w_j = \begin{cases} 1 & j \in \text{ROI} \\ 0 & \text{otherwise.} \end{cases} \tag{3.20}$$

In that case, if (3.19) is satisfied with a small threshold value γ, the resolution properties for all voxels in the ROI should be close to the desired ones specified by matrix $\boldsymbol{F}$. Note that the positive definite matrix $\boldsymbol{C}$ can be used to control the actual definition of the similarity between the target and actual PSFs. For example, one may use greater weight for the differences in the tail region.

Therefore, actual PSFs that have wider and narrower tails than those of the target PSF may get different penalties in the calculation of the similarities using (3.19) .

For an arbitrary estimator $\hat{\boldsymbol{x}}$ with a given mean-gradient matrix $\boldsymbol{G}$, its covariance matrix must satisfy [86]:

$$Cov[\hat{\boldsymbol{x}}] \geq \boldsymbol{G} \cdot \boldsymbol{J}^{-1} \cdot \boldsymbol{G}^{\boldsymbol{T}}. \tag{3.21}$$

With the weighting matrix $\boldsymbol{W}$ defined in (3.20), the minimum achievable total variance across all voxels in the ROI is given by the biased Cramer-Rao bound:

$$V_{tot} \geq tr\left\{\boldsymbol{W} \cdot Cov[\hat{\boldsymbol{x}}] \cdot \boldsymbol{W}^{\boldsymbol{T}}\right\} = tr\left\{\boldsymbol{W} \cdot \boldsymbol{G} \cdot \boldsymbol{J}^{-1} \cdot \boldsymbol{G}^{\boldsymbol{T}} \cdot \boldsymbol{W}^{\boldsymbol{T}}\right\}. \tag{3.22}$$

This equation is trivial in practice since it is specific to a very restrictive set of estimators with a given mean gradient matrix $\boldsymbol{G}$. To obtain a more useful bound, we attempt to solve the constrained minimization problem that the lowest total variance should be achieved by any estimator whose mean gradient matrix satisfies the resolution constraint (3.19):

$$V_{\min} \geq \arg\min_{s \leq \delta}(V_{tot}) = \arg\min_{s \leq \delta}\left(tr\left\{\boldsymbol{W} \cdot \boldsymbol{G} \cdot \boldsymbol{J}^{-1} \cdot \boldsymbol{G}^{\boldsymbol{T}} \cdot \boldsymbol{W}^{\boldsymbol{T}}\right\}\right). \tag{3.23}$$

Similarly as scalar MUCRB case, Karush–Kuhn–Tucker (KKT) conditions is utilized to solve this constrained minimization problem [91]. KKT multiplier λ is introduced to form:

$$\begin{aligned} f_{KKT} = & tr\left\{\boldsymbol{W} \cdot \boldsymbol{G} \cdot \boldsymbol{J}^{-1} \cdot \boldsymbol{G}^{\boldsymbol{T}} \cdot \boldsymbol{W}^{\boldsymbol{T}}\right\} \\ & + \lambda \cdot \left(tr\left\{\boldsymbol{W} \cdot (\boldsymbol{G} - \boldsymbol{F}) \cdot \boldsymbol{C} \cdot (\boldsymbol{G} - \boldsymbol{F})^{\boldsymbol{T}} \cdot \boldsymbol{W}^{\boldsymbol{T}}\right\} - \delta\right). \end{aligned} \tag{3.24}$$

Since f_{KKT} is strictly convex, the optimum mean gradient matrix $\boldsymbol{G}$ that minimizes f_{KKT} can be found by letting the derivative of (3.24) with respect to $\boldsymbol{G}$ be 0,

$$\begin{aligned} \frac{\partial}{\partial \boldsymbol{G}} f_{KKT} = & \frac{\partial}{\partial \boldsymbol{G}}\left[tr\left\{\boldsymbol{W} \cdot \boldsymbol{G} \cdot \boldsymbol{J}^{-1} \cdot \boldsymbol{G}^{\boldsymbol{T}} \cdot \boldsymbol{W}^{\boldsymbol{T}}\right\}\right. \\ & \left. + \lambda \cdot \left(tr\left\{\boldsymbol{W} \cdot (\boldsymbol{G} - \boldsymbol{F}) \cdot \boldsymbol{C} \cdot (\boldsymbol{G} - \boldsymbol{F})^{\boldsymbol{T}} \cdot \boldsymbol{W}^{\boldsymbol{T}}\right\} - \delta\right)\right] \\ = & 0. \end{aligned} \tag{3.25}$$

Given that

$$\begin{aligned} & \frac{\partial}{\partial \boldsymbol{G}}\left[tr\left\{\boldsymbol{W} \cdot \boldsymbol{G} \cdot \boldsymbol{J}^{-1} \cdot \boldsymbol{G}^{\boldsymbol{T}} \cdot \boldsymbol{W}^{\boldsymbol{T}}\right\}\right] \\ = & \boldsymbol{W}^{\boldsymbol{T}} \cdot \boldsymbol{W} \cdot \boldsymbol{G} \cdot \left(\boldsymbol{J}^{-1}\right)^{T} + \cdot \boldsymbol{W}^{\boldsymbol{T}} \cdot \boldsymbol{W} \cdot \boldsymbol{G} \cdot \boldsymbol{J}^{-1} \\ = & 2 \cdot \boldsymbol{W} \cdot \boldsymbol{G} \cdot \boldsymbol{J}^{-1}, \end{aligned} \tag{3.26}$$

and

$$\frac{\partial}{\partial \boldsymbol{G}}\left[tr\left\{\boldsymbol{W}\cdot(\boldsymbol{G}-\boldsymbol{F})\cdot\boldsymbol{C}\cdot(\boldsymbol{G}-\boldsymbol{F})^{\boldsymbol{T}}\cdot\boldsymbol{W}^{\boldsymbol{T}}\right\}\right]=2\cdot\boldsymbol{W}\cdot(\boldsymbol{G}-\boldsymbol{F})\cdot\boldsymbol{C}, \tag{3.27}$$

we could substitute (3.26) and (3.27) into (3.25) and obtain

$$\frac{\partial}{\partial \boldsymbol{G}} f_{KKT}=2\cdot\boldsymbol{W}\cdot\boldsymbol{G}\cdot\boldsymbol{J}^{-1}+2\cdot\lambda\cdot\boldsymbol{W}\cdot(\boldsymbol{G}-\boldsymbol{F})\cdot\boldsymbol{C}=0. \tag{3.28}$$

Assuming that $\boldsymbol{J}$ is non-singular, the optimum mean gradient matrix $\boldsymbol{G}$ must satisfy

$$\boldsymbol{W}\cdot\boldsymbol{G}=\boldsymbol{W}\cdot\boldsymbol{F}\cdot\left[\boldsymbol{J}+(\lambda\cdot\boldsymbol{C})^{-1}\right]^{-1}\cdot\boldsymbol{J}. \tag{3.29}$$

Note that

$$\boldsymbol{G}_{optimum}=\boldsymbol{F}\cdot\left[\boldsymbol{J}+(\lambda\cdot\boldsymbol{C})^{-1}\right]^{-1}\cdot\boldsymbol{J} \tag{3.30}$$

is a sufficient condition for (3.29) to hold.

Substitute (3.30) into (3.23), the minimum achievable total variance, with any estimator that has a mean gradient matrix satisfies (3.19), is given by [73]:

$$\begin{aligned} V_{\min} &\geq \arg\min_{s\leq\delta}(V_{tot}) \\ &= \arg\min_{s\leq\delta}\left(tr\left\{\boldsymbol{W}\cdot\boldsymbol{G}\cdot\boldsymbol{J}^{-1}\cdot\boldsymbol{G}^{\boldsymbol{T}}\cdot\boldsymbol{W}^{\boldsymbol{T}}\right\}\right) \\ &= tr\left\{\boldsymbol{W}\cdot\boldsymbol{F}\cdot\left[\boldsymbol{J}+(\lambda\cdot\boldsymbol{C})^{-1}\right]^{-1}\cdot\boldsymbol{J}\cdot\left[\boldsymbol{J}+(\lambda\cdot\boldsymbol{C})^{-1}\right]^{-1}\cdot\boldsymbol{F}^{\boldsymbol{T}}\cdot\boldsymbol{W}^{\boldsymbol{T}}\right\}. \end{aligned} \tag{3.31}$$

The vector MUCRB (3.31) incorporates a more accurate constraint of spatial resolution and allows the extension from a voxel-wise performance index to a measure of the average resolution-variance tradeoffs across an arbitrarily defined ROI. Furthermore, the minimum attainable covariance of reconstructed images can be approximated as:

$$Cov[\hat{\boldsymbol{x}}]\approx\boldsymbol{F}\cdot\left[\boldsymbol{J}+(\lambda\cdot\boldsymbol{C})^{-1}\right]^{-1}\cdot\boldsymbol{J}\cdot\left[\boldsymbol{J}+(\lambda\cdot\boldsymbol{C})^{-1}\right]^{-1}\cdot\boldsymbol{F}^{T}. \tag{3.32}$$

Similarly as scalar MUCRB, the value of λ should be very large and the resolution property offered by the efficient estimator is mainly determined by the filter function $\boldsymbol{F}$, each column of which defines the desired mean gradient function corresponding to a given voxel. In our study, these columns were usually defined as spatially shifted 3-D Gaussian functions with uniform width. By varying the shape or width of these mean gradient vectors, one can evaluate tradeoffs between the minimum attainable total variance and spatial resolution function that are given by the corresponding mean gradient matrix.

In addition, one could also compare different imaging configurations based on their performance for ROI quantification. In this case, the ROI could be regarded as a larger voxel and

covariance of any two voxels inside the ROI should be considered. If the boundary of the ROI is well-defined, the estimated total tracer-uptake in the ROI could be given by

$$tr(\hat{\boldsymbol{x}}) = \boldsymbol{z}^T \cdot \hat{\boldsymbol{x}}, \tag{3.33}$$

where $\boldsymbol{z}$ is an indicator vector for the ROI,

$$\boldsymbol{z} = \{z_j\}^T, \qquad z_j = \begin{cases} 1 & j \in \text{ROI} \\ 0 & \text{otherwise.} \end{cases} \tag{3.34}$$

Using (3.33) and (3.34), the variance and bias values associated with $tr(\hat{\boldsymbol{x}})$ can be given by [73]

$$Var_{ROI}\left[tr(\hat{\boldsymbol{x}})\right] \approx \boldsymbol{z}^T \cdot \boldsymbol{F} \cdot \left[\boldsymbol{J} + (\lambda \cdot \boldsymbol{C})^{-1}\right]^{-1} \cdot \boldsymbol{J} \cdot \left[\boldsymbol{J} + (\lambda \cdot \boldsymbol{C})^{-1}\right]^{-1} \cdot \boldsymbol{F}^T \cdot \boldsymbol{z}, \tag{3.35}$$

$$Bias_{ROI}\left[tr(\hat{\boldsymbol{x}})\right] \approx \boldsymbol{z}^T \cdot \boldsymbol{F}^T \cdot \left[\boldsymbol{J} + (\lambda \cdot \boldsymbol{C})^{-1}\right]^{-1} \cdot \boldsymbol{J} \cdot \boldsymbol{x} - \boldsymbol{z}^T \cdot \boldsymbol{x}. \tag{3.36}$$

By varying the width of the desired point spread function (column vectors of $\boldsymbol{F}$), one can derive the tradeoffs between the mean bias and the corresponding variance for ROI quantification.

3.3 Achievability of the Vector MUCRB

In general, vector MUCRB approach just provides approximations for image resolution and covariance properties and there is no guarantee that an estimator exists to achieve the variance bound. Demonstrated in [58], however, in special case that measurements obey the linear Gaussian model, the resulting post-filtered penalized weighted least-square estimator [92]

$$\begin{cases} \hat{\boldsymbol{x}}_{PWLS}(\boldsymbol{y}) = \underset{\mathbf{x}\geq 0}{\arg\max}\left[-\frac{1}{2}(\boldsymbol{y} - \boldsymbol{A}\cdot\boldsymbol{x})^T \Sigma^{-1}(\boldsymbol{y} - \boldsymbol{A}\cdot\boldsymbol{x}) - \beta \cdot R(\boldsymbol{x})\right] \\ \text{and then} \\ \hat{\boldsymbol{x}}_{PF-PWLS} = \boldsymbol{F} \cdot \hat{\boldsymbol{x}}_{PWLS}(\boldsymbol{y}) \end{cases} \tag{3.37}$$

could exactly achieve the bound (3.31), with the quadratic penalty function satisfying

$$\beta \cdot \nabla^2 R(\boldsymbol{x}) = (\lambda \cdot C)^{-1}, \tag{3.38}$$

where Σ^{-1} is the inverse of the covariance matrix of $\boldsymbol{y}$, which could be considered as a diagonal matrix for independent measurements.

As previously discussed, the estimation of image in emission tomography in linear Poisson model could be commonly achieved by the post-filtered penalized maximum-likelihood (PF-PML) algorithm, given by:

$$\begin{cases} \hat{\boldsymbol{x}}_{PML}(\boldsymbol{y}) = \underset{\mathbf{x}\geq 0}{\arg\max}\left[L(\boldsymbol{x},\boldsymbol{y}) - \beta \cdot R(\boldsymbol{x})\right] \\ \text{and then} \\ \hat{\boldsymbol{x}}_{PF-PML} = \boldsymbol{F} \cdot \hat{\boldsymbol{x}}_{PML}(\boldsymbol{y}). \end{cases} \tag{3.39}$$

Intuitively this estimator may also achieve the vector MUCRB in high-counts detection situation under which the linear Poisson model could be simplified as the linear Gaussian model. Actually it is demonstrated in [58, 91] that the PF-PML estimator with penalty function defined in (3.38) is an efficient estimator that asymptotically achieves the vector MUCRB with the mean gradient vector defined in (3.30) in high count-rate situation. Normally the total counts of detected photons could be adjusted by some system parameters, such as activities of tracer solution to be injected to the object, and the total imaging time. In most high resolution imaging applications, the requirement for photon counts could be satisfied and therefore it is meaningful to implement vector MUCRB approach for quantifying the fundamental system performance.

For a given physical system design and a deterministic image function, (3.30) indicates that two factors affect the spatial resolution function of the efficient estimator: regularization parameter β and the filter function $\boldsymbol{F}$. Normally one should restrict regularization factor (β=1/λ) to a small value to satisfy resolution constraint (3.30), and thus the regularization in (3.39) becomes so tiny in the reconstruction that *a priori* information is not important any more. The resolution property offered by the efficient estimator is mainly determined by the filter function $\boldsymbol{F}$.

The required computation load for evaluating the resolution-variance tradeoff for one voxel or bias-variance tradeoff for ROI quantification is similar as one reconstruction [73]. Given that the construction of an average resolution-variance curve typically requires less than ten points in the object-space, the formulations derived allow quantitative system evaluations based on the statistical image quality with a reasonable computation time. In comparison, a large number of reconstructions (usually more than 100) are required in brute-force MC simulation to achieve the variance with sufficient accuracy. Therefore, vector MUCRB can be used as an analytical performance index for comparing different SPECT system and thus be considered as the first step towards the system optimization.

3.4 Monte Carlo Simulation Results

3.4.1 Introduction to a Simulated SPECT System and Phantom

The vector MUCRB approach was implemented in the design of a MRI compatible SPECT system [93], based on the energy-resolved photon counting (ERPC) semiconductor detector [36]. In this system six detectors were used to form a stationary ring, shown in Figure 3.4. All detectors were assumed to be 100% efficient for detecting I-125 gamma rays. The depth-of-interaction (DOI) effect was not modeled in this simulation. In the system design, four types of detectors were compared, which have the similar active area of roughly 4.5 cm × 4.5 cm but different pixelation configurations (44 × 44, 90 × 90, 128 × 128 and 220 × 220 square pixels of 1 mm × 1 mm, 0.5 mm × 0.5 mm, 0.35 mm × 0.35 mm and 0.2 mm × 0.2 mm in size respectively). The collimator system consists of 6 apertures that are made of tungsten sheet of 5 mm in thickness. In

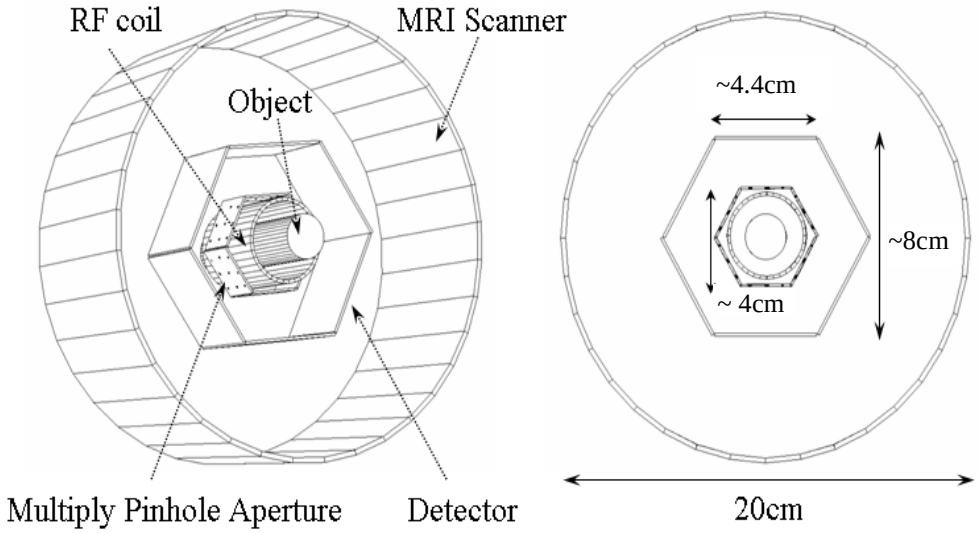

Figure 3.4 The proposed MRI-compatible SPECT system. Left penal: schematic of the SPECT system inside a MRI scanner. Right panel: dimensions of different components

the aperture design, we also compare four pinhole configurations (4×4, 5×5, 6×6 and 7×7) with pinhole diameter of 200 μm. As a basic design rule, the pinholes locations were so chosen that (a) the overlapping between projections through multiple pinholes was minimized and (b) the entire detector area are fully utilized by projections. Design parameters of several apertures are given in

Table 3.1. For these configurations, the projections of the resolution phantom on the detector are compared in Figure 3.5. All pinholes have sharp knife-edges and acceptance cones of 90 degrees on both sides. Photon scattering and penetration in aperture were not modeled. This

simulation study was carried out using a Monte Carlo package that we developed for the various SPECT applications [31, 42, 74].

Table 3.1 Design Parameters of Four Aperture Configurations

No. of pinholes	16	25	36	49
Center-to-aper. dist.	20mm	20mm	21mm	23mm
Det.-to-aper. Dist.	27mm	22mm	19mm	17mm
Avg. pinhole dist.	4.68mm	4.19mm	3.85mm	3.61mm

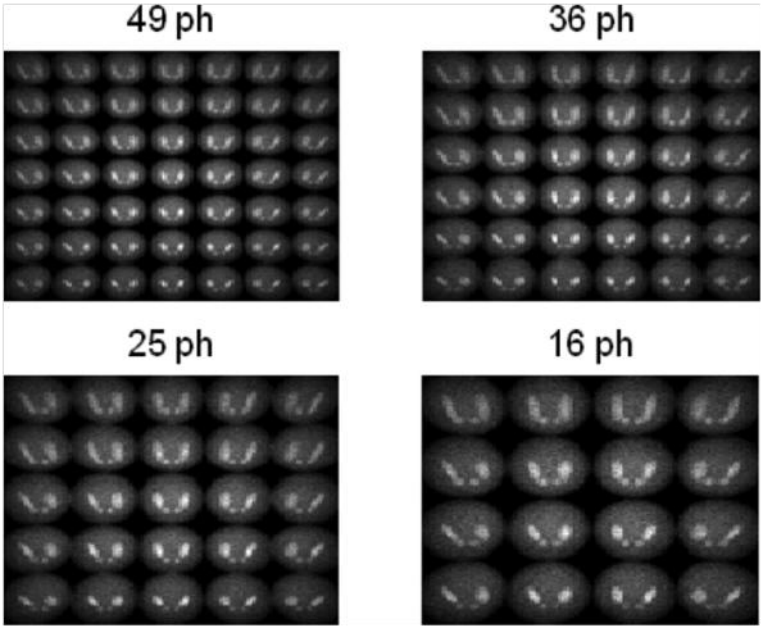

Figure 3.5 Projections of the resolution phantom on the detector area

The imaging study was based on a simulated phantom that has a cylindrical volume of 1 cm long and 1 cm in diameter. The object-space was divided into 64 × 64 × 64 voxels. Each voxel was 200 μm × 200 μm × 200 μm in size. The source contains a total activity of 100 μCi. The phantom consists of two hot spherical volume of 1 mm in diameter and two hot ellipsoidal volumes that have half-axes of 0.4 mm, 0.4 mm and 0.8 mm respectively. These hot features are superimposed on a continuous background (A_B) with activity concentration of A_{H1}. Several small hot and cold spheres are inserted in the two hot ellipsoidal volumes. The activity concentrations in these small spheres are A_{HR} and A_{CR} respectively. The ratios of the activity concentrations

within different regions in the simulated phantom are A_{H1}: A_B: A_{HR}: A_{CR} =10:1:20:1. For estimating the average variance over multiple voxels, we selected nine control points in the center of the phantom. A cross section of the phantom is shown in Figure 3.6.

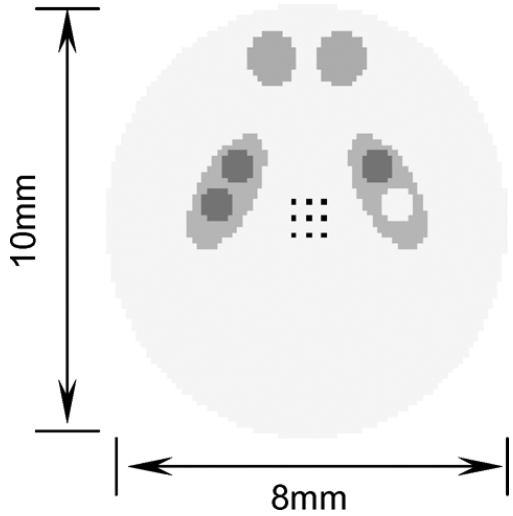

Figure 3.6 Cross sections of the resolution phantom. Nine control points were selected (as shown with black dots) for the calculation of the average resolution-variance tradeoffs

3.4.2 Validation of vector MUCRB Calculation

For linear Poisson problems, the vector MUCRB (3.31) can be asymptotically achieved by post-filtered penalized maximum likelihood (PF-PML) estimator. In this study, we compared the MUCRB with the empirical variance achieved with PF-PML estimators for the voxel in the center of the phantom. To ensure the convergence of the PML estimators, 250 iterations were used in reconstructions. The total observation time is 1 hour. The empirical variance was obtained from 100 random realizations. In this comparison, a reasonably good agreement between the MUCRB and the empirical variance was demonstrated in Figure 3.7. Since the derivation of vector MUCRB does not include the non-negativity constraint that is enforced in PML estimators [58, 72], the minimum attainable variance predicted by vector MUCRB tends to be greater than that achieved by actual PM-PML estimators in low count-rate situations. However, the difference between vector MUCRB and empirical variance of PF-PML estimator was kept below 7%, even with a small regularization factor.

In this study, we also compared the mean gradient vectors predicted by (3.30) for the center voxel and local impulse responses obtained from corresponding PF-PML estimators. An excellent agreement is demonstrated with the example shown in Figure 3.8, which indicates that it is reasonable to use mean gradient vector instead of local impulse response to quantify the system resolution property.

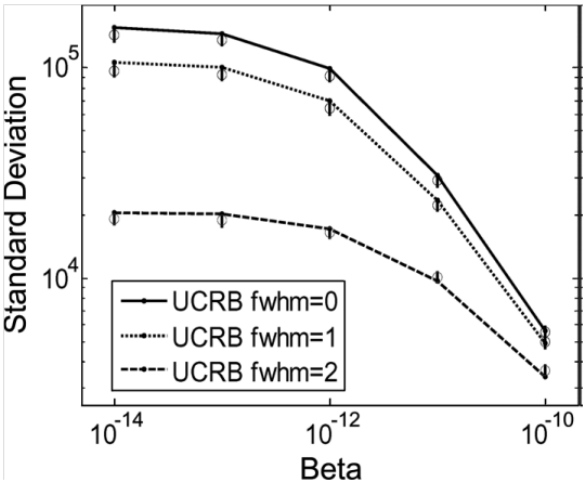

Figure 3.7 Comparing the MUCRB (curves) and the empirical variance value (circles with error bars) derived from Monte Carlo simulation. The standard deviation values were derived for the center pixel.

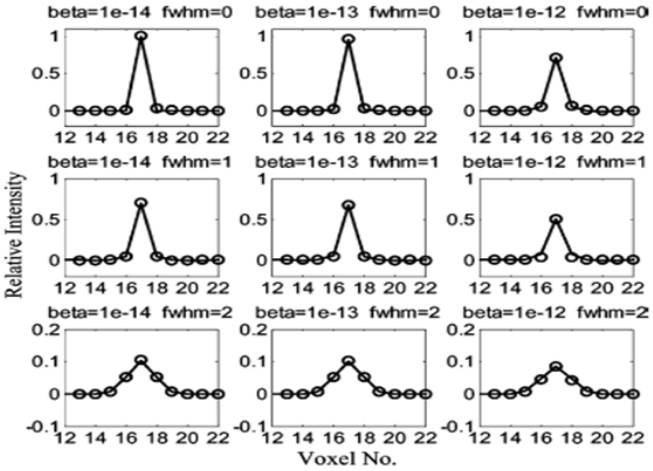

Figure 3.8 Solid lines: mean gradient vectors of the optimum estimator that achieve the MUCRB. Circles: local impulse response obtained with actual PF-PML estimators.

3.4.3 Aperture Design Based on Resolution-Variance Tradeoffs Derived by Vector MUCRB

In this study, we compared the imaging performances of the SPECT system with four different multiple-pinhole apertures. The smaller number of pinhole could provide larger magnification and thus better resolution property, at the cost of the lower sensitivity and poorer statistical performance. Therefore, it is not straightforward to find the best aperture design for imaging the simulated phantom. The average resolution-variance tradeoffs over center nine voxels derived for different aperture configurations are compared in Figure 3.9. It has shown that the best imaging performance was achieved with the 4×4 pinhole aperture. This conclusion was validated by the comparison of reconstructions in Figure 3.10. These images were chosen to have similar average variance over the control points. Clearly imaging system with 4×4 pinhole aperture was able to provide the optimum resolution in resolving the features in reconstructed images. The simulation results demonstrate that the vector MUCRB can be used to evaluate the fundamental tradeoffs between resolution and image variance offered by a given system design. This method may serve as the basis for evaluating imaging system design and system optimization.

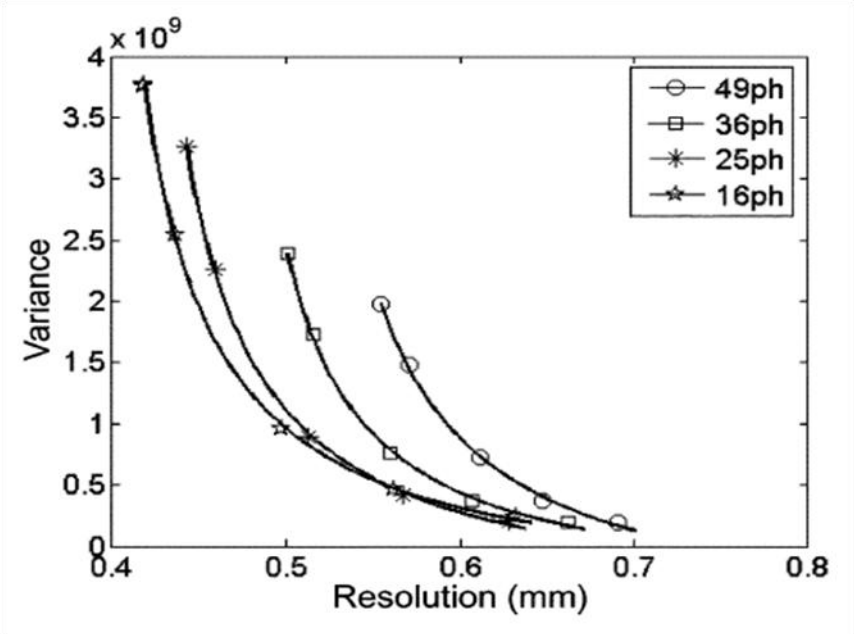

Figure 3.9 Average resolution-variance tradeoffs achieved with four system configurations as detailed in

Table 3.1

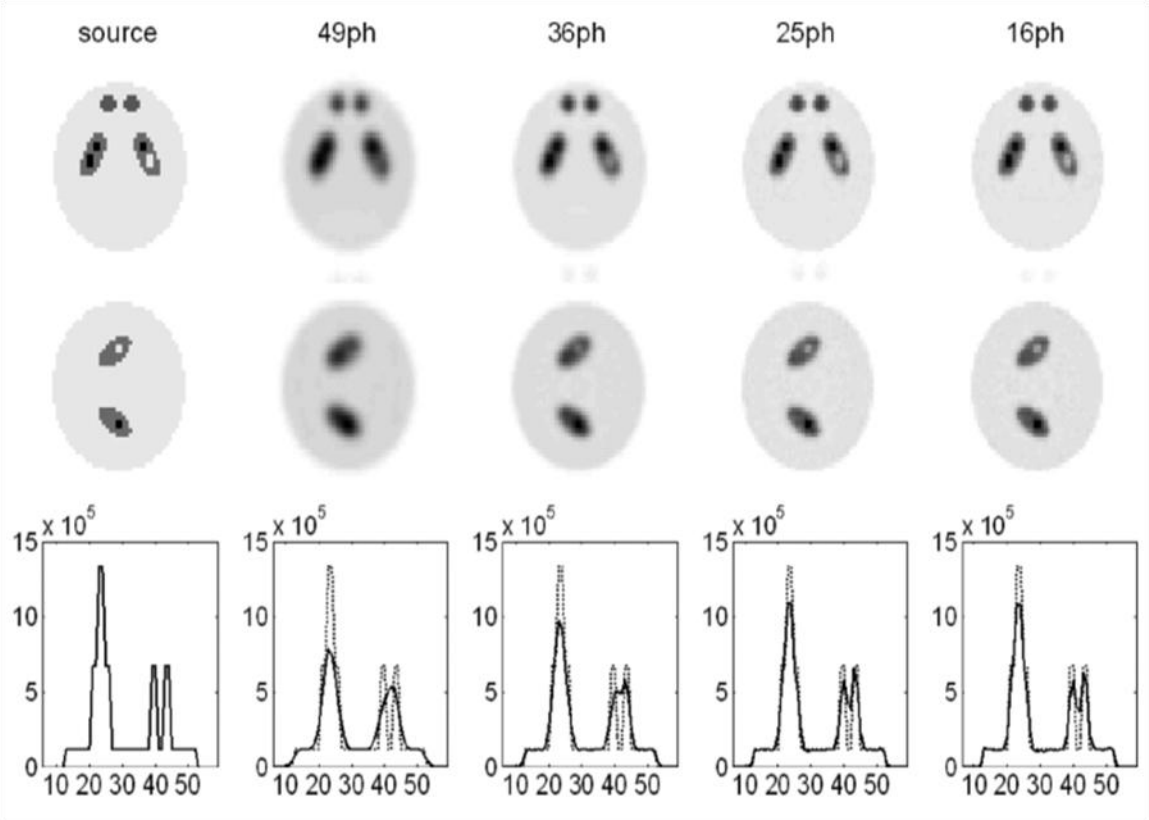

Figure 3.10 Reconstructed images of the resolution phantom. The average variances over the nine control-points are similar for all four cases compared.

3.4.4 Detector Pixelation Comparison Based on Bias-Variance Tradeoffs Derived by Vector MUCRB

We have also used the proposed SPECT system to evaluate the bias-variance curves for ROI quantification, derived by (3.35) and (3.36). In this case, the aperture with 6×6 pinhole was used and the detector pixel size was changing from 1 mm × 1 mm to 0.2 mm × 0.2 mm. A ROI was defined at the center of the FOV, which has a spherical volume of 1.6 mm in diameter. The bias and variance values are plotted in Figure 3.11. As demonstrated in this figure, improving detector resolution from 1 mm × 1 mm to 0.5 mm × 0.5 mm led to a significant improvement in ROI quantification. Further improved resolution offered only a modest performance gain. These indications were confirmed by reconstructed images shown in Figure 3.12, which are corresponding to the four operation-points in Figure 3.11. Once again, consistent results have been demonstrated between the analytically-derived system performance and the visual appearance of reconstructed images.

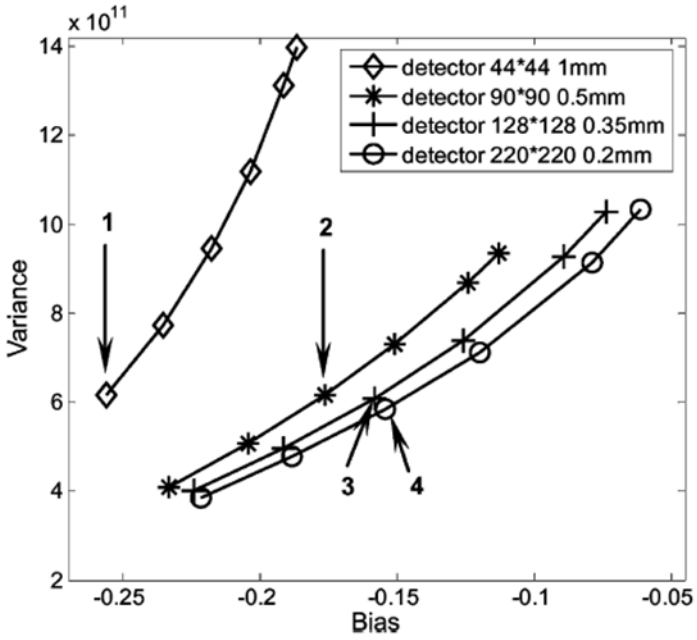

Figure 3.11 Bias-variance curves derived for four detector pixelations.

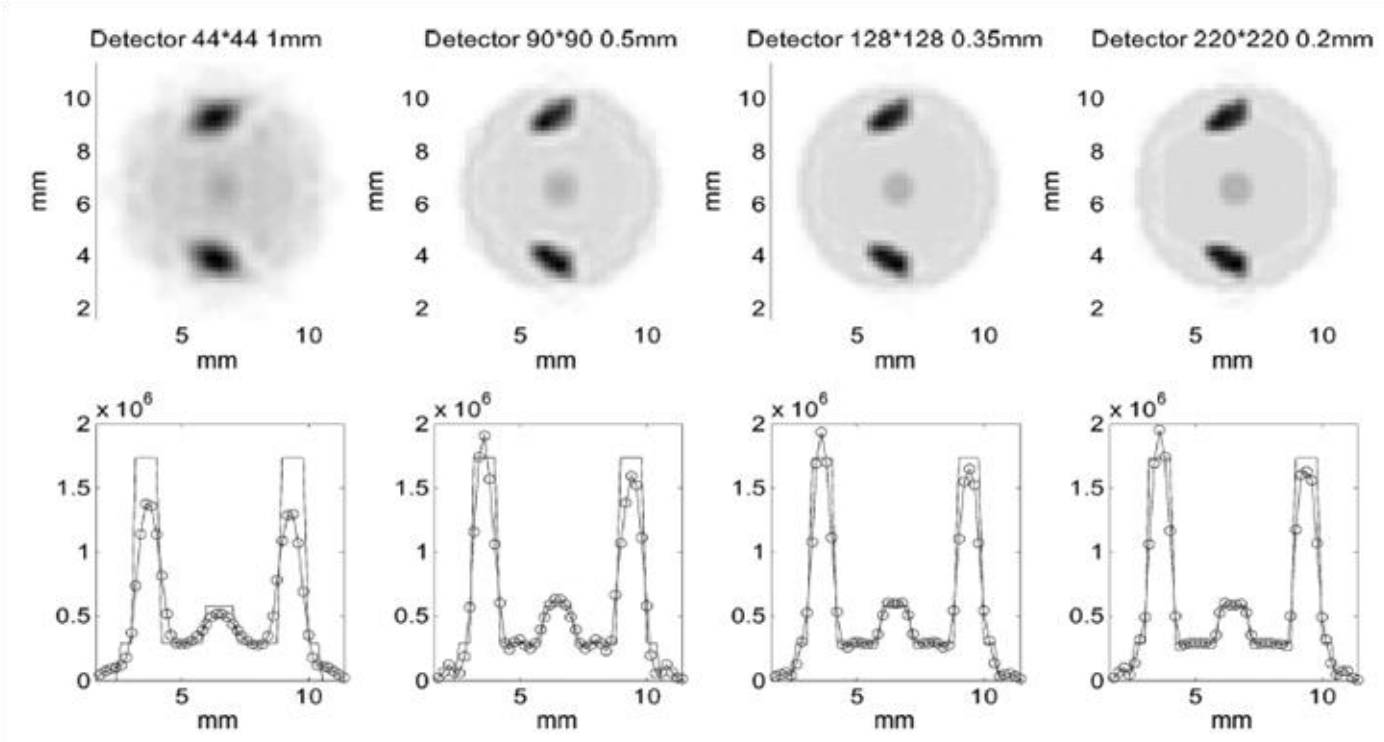

Figure 3.12 Reconstructed images of the phantom with a hot blob at the center. These figures serve as a visual confirmation of the bias-variance tradeoff curves. The four columns are corresponding to the operation points highlighted in Figure 3.11. The pixel sizes simulated and the total number of pixels in each detector are given on the top.

3.5 Summary

In this work, a vector modified uniform Cramer-Rao bound (MUCRB) is presented. It gives the minimum attainable total (or average) variance of an unknown deterministic vector parameter, given that the resultant mean gradient matrix satisfies a certain constraint. The vector MUCRB can be used to evaluate and compare the fundamental resolution-variance tradeoffs or bias-variance tradeoffs offered by a given system design. For linear Poisson model that commonly used to describe SPECT systems, the optimum resolution-variance tradeoffs predicted by vector MUCRB can be achieved asymptotically by post-filtered penalized maximum likelihood estimators with well-defined penalty function. This ensures that the system performance predicted is meaningful in routine practice.

The use of the vector MUCRB for evaluating system designs was demonstrated in a series of MC simulations with a simulated stationary SPECT system. The resolution-variance tradeoffs and bias-variance tradeoffs, predicted by the MUCRB approach, are consistent with that obtained from reconstructed images. The evaluation of single point intensity or ROI quantification on the tradeoff curve requires a computation load similar to that of a single image reconstruction. Given that the construction of a resolution-variance curve typically requires less than ten points in the resolution-variance space, the formulations derived allow quantitative system evaluations base on the statistical image quality with a reasonable computation time. This method may serve as the basis for SPECT imaging system performance optimization.

CHAPTER 4

NON-UNIFORM OBJECT-SPACE PIXELATION (NUOP) APPROACH

As introduced before, one of the recent trends in SPECT instrumentations is to achieve an ultra-high resolution for imaging small lab animals. Several recent developments include the SemiSPECT reported by Kastis et al. [43], the SiliSPECT under development by Peterson et al. [44], the MediSPECT proposed (and evaluated) by Accorsi et al. [39] and the U-SPECT-II proposed by Beekman et al. [34]. We have recently developed a prototype single photon emission microscope (SPEM) system for mouse brain studies [40-42]. This system was based on the intensified EMCCD cameras that offer an excellent intrinsic resolution for low energy gamma rays. It was demonstrated that the current dual-headed SPEM system is capable of visualizing a very small number of radiolabeled T cells in mouse brain.

For most of reconstruction problems, image functions are typically represented with equally sized cubic voxels throughout the object space. Despite its simplicity, the uniform sampling provided by the cubic voxels could be less efficient for the high resolution imaging application. In order to fully utilize the excellent resolution capability, the reconstruction of images requires the use of very small voxel sizes. If such voxel sizes are used uniformly throughout the entire object-space, the reconstruction would involve a tremendous amount of computation. For instance, despite the extensive effort in adapting the reconstruction task in parallel computing environment, a single SPEM reconstruction typically requires tens of hours to a few days to complete. Furthermore, the biggest disadvantage of high resolution SPECT instrumentations is the limited sensitivity and relatively poor statistical property. So we intend to optimize the system configuration for maximizing system efficiency for collecting useful information about the object. A practical optimization process needs to rapidly evaluate system performance, which could be expressed as tradeoffs between system resolution and statistical noise characteristic in our study. Even with vector MUCRB to analytically quantify the tradeoffs, instead of time-consuming brute-force MC simulation, the required computation load, equivalent with multiple reconstructions, is still very huge. In conclusion, rapid image reconstruction process and system performance evaluation are crucial for practical applications and system optimization of the high resolution imaging system.

Since the primary task of a high-resolution imaging system is to reveal the microscopic structure inside a target-region, one may tolerate lower spatial resolutions in areas outside the

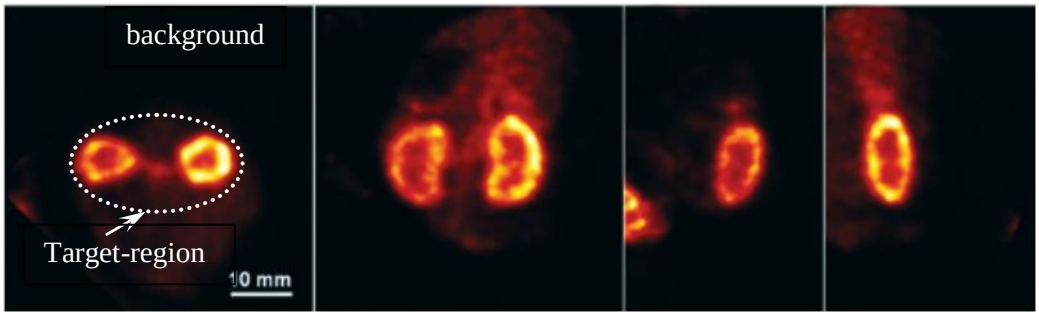

Figure 4.1 Mouse kidney scan. Intensity slices through the volumetric reconstruction of the mouse kidneys, labeled with [^{99m}Tc]: (A) transverse plane; (B) coronal plane; and (C1 and C2) sagittal plane. The ROI shown only takes up a small part of object-space

target-region. In many cases, the interested features seize only a small part of the object-space, and others are background or indifferent features, as seen in Figure 4.1 [41]. Therefore, larger voxels could be used outside the target-region for improved computation efficiency, compared to the conventional uniform object-space pixelation scheme.

In this section a non-uniform object-space pixelation (NUOP) approach is proposed to divide the object into smaller voxels for target-regions, and into larger voxels in areas that are relatively smooth and/or less important to the reconstruction of the target-regions. This method does not require any *a priori* information of the object to be imaged. Instead, it is determined adaptively according to (a) the imaging information being acquired by the detection system, (b) the known system response function and (c) the user-defined target-region. The objective of the NUOP approach is to offer a significant reduction of computation-effort without sacrificing the imaging quality in the target-region. This makes it particularly useful for high resolution imaging application. In addition, this development also allows rapid evaluation of the first and second order statistics of reconstruction images using analytical approximations. This provides crucial information for the derivation of several statistical system performance indices that could be used in system design and optimization.

4.1 Non-Uniform Object-Space Pixelation

The final goal of the high resolution imaging study is to obtain focused microscopic images of a local target-region. It is possible to speed up the reconstruction with a non-uniform object-space pixelation approach as detailed below [94]:

- **Step 1**: The object-space is modeled with sufficiently small and uniformly-sized voxels, originally represented by $\boldsymbol{x}=[x_1, x_2, \ldots, x_N]^T$. For example, one could use voxel-sizes smaller than or equal to 1/2 of the expected system resolution. This step is performed only once for a

given system and the resultant large and comprehensive system response functions are stored for future use.

- **Step 2**: A target-region will be specified based on some initial knowledge, such as a preliminary image of the object. This image could be obtained in a short study using a uniform imaging time distribution, and/or apertures having large pinholes to ensure a good counting statistics, and relatively coarse and equally-sized voxels in the object-space for fast reconstruction [47]. The image obtained could be used to identify the target-region and to provide a rough estimate of the underlying source function. In addition, initial knowledge may also be achieved from other imaging modalities, e.g., computed tomography (CT), positron emission tomography (PET), and magnetic resonance imaging (MRI).
- **Step 3**: According to the target-region chosen in step 2, a non-uniform object-space pixelation scheme could be defined by a pixelation density function (PDF), whose integer element, called D-value, specifies the pixelation density for the related source voxel in the proposed NUOP approach. Several systematic approaches for defining pixelation density functions will be discussed later. Here, we describe a simple algorithm for the actual rebinning process with a given pixelation density vector ***D***. The entire object-space is

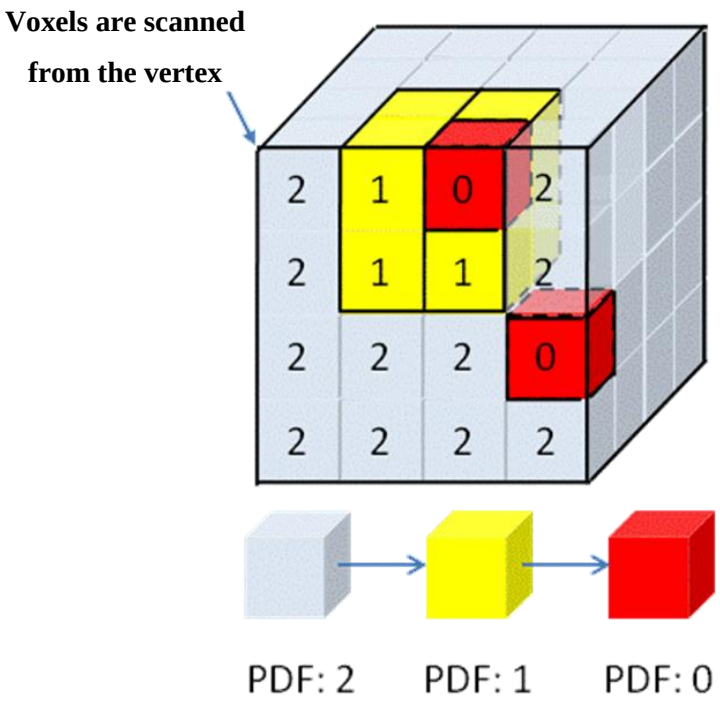

Figure 4.2 A simple example to illustrate the rebinning process. The PDF function is defined with different colors. In this case, 4×4×4 object-space is divided into one blue largest voxel, two yellow voxels and two original red voxels.

scanned in descending order of the D-value. For a specified D-value of D, $2^D \times 2^D \times 2^D$ adjacent voxels from a certain vertex of the object-space start to be scanned, and only those with the same D-value will be combined together and the rest are ignored. Once the rebinning process for the current voxel cube is completed, the algorithm goes on to scan the next $2^D \times 2^D \times 2^D$ voxel cube and rebin voxels in the same way. After all the voxels with the same D-value are processed, the algorithm will then move on to scan object-space with D-value of D-1 and keep doing so until all voxels with D-values greater than 0 are processed. For a given source voxel, the corresponding D-value of 0 indicates that the voxel does not need to be rebinned. The voxel numbers after rebinning are given based on the certain order, by which they are formed in the rebinning process. In practice, the user has complete freedom in defining the rebinning parameters, such as D-value and the corresponding subset of voxels to be combined, in response to the specific imaging task. A simple example is shown in Figure 4.2 to illustrate this rebinning process.

The object-space, after this rebinning process, is represented by a new vector $\boldsymbol{x}_{Rebin}=[(x_{Rebin})_1,(x_{Rebin})_2,\ldots,(x_{Rebin})_P]^T$, which would typically have a much smaller number of voxels. To simplify the discussion, this rebinning process could be represented by a transformation matrix $\boldsymbol{S}$ ($N \times P$ in size) as follows:

$$\boldsymbol{S}_{ii'} = \begin{cases} 1 & \text{if voxel } i \text{ in the uniform object} \\ & \text{space is rebinned into voxel } i' \\ & \text{in the non-uniform object space,} \\ 0 & \text{otherwise.} \end{cases} \tag{4.1}$$

Note that this transform matrix has a one-to-one correspondence with pixelation density function $\boldsymbol{D}$. With the NUOP approach, certain voxels originally in the uniformly pixelated object-space can be combined to form larger voxels.

- **Step 4**: A new system response function (SRF) $\boldsymbol{A}_{Rebin}$ is needed to reflect the transformation from the non-uniformly pixelated object-space to the detector space, which is given by

$$\boldsymbol{A}_{Rebin} = \boldsymbol{A} \cdot \boldsymbol{S}. \tag{4.2}$$

In this matrix, element $\left(\boldsymbol{A}_{Rebin}\right)_{mp}$ gives the probability of a gamma-ray emitted from the p'th rebinned source voxel and detected by the m'th detector pixel, which is equal to the sum of probabilities related to all the original source voxels to be combined together. The required computation load for generating $\boldsymbol{A}_{Rebin}$ is much smaller than that for generating comprehensive system response function $\boldsymbol{A}$. The expectation of the measurement is related to the underlying source vector $\boldsymbol{x}_{Rebin}$ as:

$$\bar{\boldsymbol{y}} = \boldsymbol{A}_{Rebin} \cdot \boldsymbol{x}_{Rebin} + \boldsymbol{r}. \tag{4.3}$$

The new source function $\boldsymbol{x}_{Rebin}$ then represents the radioactive concentration contained in each voxel in the non-uniformly pixelated object-space.

- **Step 5**: Subsequently, the underlying source function can be reconstructed using the penalized maximum-likelihood algorithm,

$$\left(\hat{\boldsymbol{x}}_{Rebin}\right)_{PML} = \underset{x_{Rebin} \geq 0}{\arg\max} \quad \left[L_{Rebin}(\boldsymbol{x}_{Rebin}, \boldsymbol{y}) - \beta \cdot R_{Rebin}(\boldsymbol{x}_{Rebin})\right], \tag{4.4}$$

where the corresponding log-likelihood function $L_{Rebin}(\boldsymbol{x}_{Rebin}, \boldsymbol{y})$ is given by

$$L_{Rebin}(\boldsymbol{x}_{Rebin}, \boldsymbol{y}) = \sum_{m=1}^{M} \{y_m \log[\bar{y}_m(\boldsymbol{x}_{Rebin})] - [\bar{y}_m(\boldsymbol{x}_{Rebin})]\}. \tag{4.5}$$

In this study, we used a quadratic roughness penalty matrix $\boldsymbol{R}_{Rebin}(\boldsymbol{x}_{Rebin})$ in the non-uniformly pixelated object-space, which is defined as:

$$\boldsymbol{R}_{Rebin} = \boldsymbol{S}^T \cdot \boldsymbol{R} \cdot \boldsymbol{S}, \tag{4.6}$$

where matrix $\boldsymbol{R}$ is the quadratic roughness penalty matrix in the uniformly pixelated object-space. According to the definition, if two source functions, represented in uniformly and non-uniformly pixelated object-spaces, are related by:

$$\hat{\boldsymbol{x}} = \boldsymbol{S} \cdot \hat{\boldsymbol{x}}_{Rebin}. \tag{4.7}$$

The corresponding penalty functions would be identical,

$$\begin{aligned} R_{Rebin}(\hat{\boldsymbol{x}}_{Rebin}) &= \hat{\boldsymbol{x}}_{Rebin}{}^T \cdot \boldsymbol{R}_{Rebin} \cdot \hat{\boldsymbol{x}}_{Rebin} \\ &= \hat{\boldsymbol{x}}_{Rebin}{}^T \cdot \boldsymbol{S}^T \cdot \boldsymbol{R} \cdot \boldsymbol{S} \cdot \hat{\boldsymbol{x}}_{Rebin} = \hat{\boldsymbol{x}}^T \cdot \boldsymbol{R} \cdot \hat{\boldsymbol{x}} = R(\hat{\boldsymbol{x}}). \end{aligned} \tag{4.8}$$

- **Step 6**: For convenience, the estimator, $\left(\hat{\boldsymbol{x}}_{Rebin}\right)_{PML}$, could be restored back into the uniform object-space by the following operation

$$\hat{\boldsymbol{x}}_{Restore} = \boldsymbol{S} \cdot \left(\hat{\boldsymbol{x}}_{Rebin}\right)_{PML}. \tag{4.9}$$

This restoration process simply assigns the value of each rebinned voxel back to those voxels that were originally in the uniform object-space, but combined into corresponding rebinned voxel during the rebinning process. Finally, we could also filter the resultant estimator, $\hat{\boldsymbol{x}}_{Restore}$, to obtain a smoother reconstruction:

$$\hat{\boldsymbol{x}}_{PF-Restore} = \boldsymbol{F} \cdot \hat{\boldsymbol{x}}_{Restore} = \boldsymbol{F} \cdot \boldsymbol{S} \cdot \left(\hat{\boldsymbol{x}}_{Rebin}\right)_{PML}. \tag{4.10}$$

4.2 Determination of the Rebinning Strategy

The key to the non-uniform object-space pixelation approach is to find a meaningful rebinning strategy. In order to reduce the amount of computation involved, one would use as less source voxels as possible. However, a discrete image representation with a sufficiently fine sampling is necessary to ensure that the subtle details in the ROI are retained in reconstruction, which is particularly important for ultra-high resolution imaging applications. In our study, we propose an adaptive procedure for defining the spatially variant pixelation density function.

At first, the target-region (TR) contains all the features we are interested in. It is natural that the finest voxel should be kept in this region to remain the image quality of the required features.

Second, the partial volume effect should also be taken into account. Since multiple voxels in the uniform object-space can be represented by a lager voxel with a single intensity value in the non-uniform object-space, artifacts must be inevitably introduced in reconstruction if activities contained in these voxels suffer from dramatic changes, e.g., voxels are distributed along a sharp edge. For safety purpose, voxels in the regions with sharp features (SR) do not need to be rebinned in order to reserve corresponding information in the reconstruction.

In addition to TR or SR, rebinning strategy in other areas could be determined by whether voxels are important to image quality inside the ROI or not. For SPECT imaging problem, the measured projection is characterized by a set of independent Poisson variables, so the element of the Fisher information matrix is given by

$$\boldsymbol{J}_{\boldsymbol{ij}} = -E\left\{\left[\frac{\partial^2}{\partial x_i \partial x_j} L(\boldsymbol{x}, \boldsymbol{y})\right]\right\} = \sum_m \frac{a_{mi} a_{mj}}{\bar{y}_m}, \tag{4.11}$$

where a element a_{mn} in the system response function $\boldsymbol{A}$ gives the probability of a gamma-ray emitted from the *n*'th source voxel and detected by the *m*'th detector pixel. $\bar{y}_m$ is the expectation of measurement at the *m*'th detector pixel. It is easily seen that each column of Fisher Information matrix measures the weighted correlation between the responses of the system to gamma ray emission at a given pixel of interest and at all other pixels in object-space. As previously demonstrated by Qi et al. [64, 95], the reconstructed image property (such as impulse response function and local covariance structure) at or around a given voxel is affected mostly by those voxels that have relatively large (weighted) correlation with the voxel-of-interest. Based on this concept, for a given target-region, one could combine columns of Fisher information matrix corresponding to the voxels inside the target-region to generate a new vector $\boldsymbol{Q}$:

$$\boldsymbol{Q} = \sum_{k \in TR} \boldsymbol{J} \cdot \boldsymbol{e}_k, \tag{4.12}$$

where $\boldsymbol{e}_k$ is the k'th unit vector. The vector $\boldsymbol{Q}$ reveals the "relative importance" between each source voxel and the target-region [94]. With this vector, one could use a finer pixelation for the target region and other correlated regions, and a coarser pixelation for regions less correlated to the target-region. This would not lead to significant degradation in reconstructed images of the target-region. The corresponding pixelation density function $\boldsymbol{D}$ is therefore defined as a function of $\boldsymbol{Q}$.

At last, the rebinning strategy may also be determined by the distance between each voxel and the center of the target-region. It is straightforward to remain fine sampling rate in the region close to the target-region

Considering these four factors, a pixelation density function (PDF) $\boldsymbol{D}$ is defined below, whose integer element, called D-value, specifies the pixelation density for the related source voxel.

$$D_k = \begin{cases} 0 & k \in \text{TR or SR} \\ \text{int}\left\{ \log_s \dfrac{\max\left[\left(Q_j\right)^{u1} \cdot \left(1 / r_j^t \right)^{u2}, j = 1, \cdots, N \right]}{\left(Q_k\right)^{u1} \cdot \left(1 / r_k^t \right)^{u2}} \right\} & \text{otherwise} \end{cases} \quad k = 1, \cdots, N, \tag{4.13}$$

where operation int[α] returns the maximum integer that is smaller or equal to α. This equation assigns the ranking of k'th source voxel in the original object space depending on its correlation to voxels inside the target-region Q_k and distance to the center of the target-region r_k. Constant t controls the degree of the distance-dependency. Constant s (>1) controls the rate, by which D_k changes with Q_k and r_k. In general, small D-value would be assigned to voxels with large correlation and/or short distance to the target-region, which indicates that relatively fine sampling is employed in the corresponding region. The minimum D-value is zero and the related voxels should not be rebinned.

In the simulation later, three types of rebinning schemes will be studied. At first, we set $u1$=1 and $u2$=0, and pixelation density function thus depends only on the relative importance between each source voxel and the target-region, which could be derived by Fisher information matrix. In addition, we set $u1$=1 and $u2$=1, and pixelation density function depends on both FIM and how far a voxel is from the center of a target-region, which may give more aggressive reduction of the number of unknowns in the object space. At last, we set $u1$=0 and $u2$=1, and the pixelation density function depends only on the linear distance $\boldsymbol{r}$. The resultant rebinning scheme

emphasizes the importance of the target-region only, without considering the correlation with other regions in the object-space.

4.3 Vector MUCRB in the Non-Uniformly Pixilated Object-Space

The NUOP approach could be adapted to the vector modified Uniform Cramer-Rao bound for the evaluation of the image quality in the reconstruction. Similar to the derivation in the uniformly pixelated object-space [72, 73], the element of Fisher information matrix in non-uniform object-space is given by [96]:

$$\left(\boldsymbol{J}_{\boldsymbol{Rebin}}\right)_{ij} = -E\left\{\left[\frac{\partial^2}{\partial\left(x_{\mathrm{Re}bin}\right)_i \partial\left(x_{\mathrm{Re}bin}\right)_j} L_{\mathrm{Re}bin}(\boldsymbol{x}_{\mathrm{Re}bin}, \boldsymbol{y})\right]\right\}. \tag{4.14}$$

Substitute the log-likelihood function $L_{Rebin}(\boldsymbol{x}_{Rebin}, \boldsymbol{y})$ in the non-uniformly pixelated object-space (4.5) into (4.14) and we have the corresponding Fisher information matrix in the matrix form as:

$$\boldsymbol{J}_{Rebin} = \left(\boldsymbol{A}_{Rebin}\right)^T \cdot \boldsymbol{diag}\left\{\bar{y}_m\right\} \cdot \boldsymbol{A}_{Rebin} = \left(\boldsymbol{A} \cdot \boldsymbol{S}\right)^T \cdot \boldsymbol{diag}\left\{\bar{y}_m\right\} \cdot \left(\boldsymbol{A} \cdot \boldsymbol{S}\right). \tag{4.15}$$

Similar to (3.32), we have the corresponding covariance matrix for $\left(\hat{\boldsymbol{x}}_{Rebin}\right)_{PML}$ approximately given by

$$Cov\left[\left(\hat{\boldsymbol{x}}_{Rebin}\right)_{PML}\right] \approx \left[\boldsymbol{J}_{Rebin} + \beta \cdot \boldsymbol{R}_{Rebin}\right]^{-1} \cdot \boldsymbol{J}_{Rebin} \cdot \left[\boldsymbol{J}_{Rebin} + \beta \cdot \boldsymbol{R}_{Rebin}\right]^{-1}, \tag{4.16}$$

and the mean gradient matrix of the corresponding reconstruction could be approximated as:

$$\boldsymbol{G}\left(\boldsymbol{x}_{Rebin}\right) = \frac{\partial E\left(\hat{\boldsymbol{x}}_{Rebin}\right)}{\partial \boldsymbol{x}_{Rebin}} \approx \left[\boldsymbol{J}_{Rebin} + \beta \cdot \boldsymbol{R}_{Rebin}\right]^{-1} \cdot \boldsymbol{J}_{Rebin}. \tag{4.17}$$

Since both post-filtering and restoring are simple linear operations, we can approximate the covariance matrix and mean gradient matrix for $\hat{\boldsymbol{x}}_{PF-Restore}$ as [50]:

$$\begin{aligned} &Cov\left[\hat{\boldsymbol{x}}_{PF-Restore}\right] \approx \boldsymbol{F} \cdot \boldsymbol{S} \cdot Cov\{\left(\hat{\boldsymbol{x}}_{Rebin}\right)_{PML}\} \cdot \boldsymbol{S}^T \cdot \boldsymbol{F}^T \\ &= \boldsymbol{F} \cdot \boldsymbol{S} \cdot \left[\boldsymbol{J}_{Rebin} + \beta \cdot \boldsymbol{R}_{Rebin}\right]^{-1} \cdot \boldsymbol{J}_{Rebin} \cdot \left[\boldsymbol{J}_{Rebin} + \beta \cdot \boldsymbol{R}_{Rebin}\right]^{-1} \cdot \boldsymbol{S}^T \cdot \boldsymbol{F}^T, \end{aligned} \tag{4.18}$$

and

$$\boldsymbol{G}\left(\boldsymbol{x}_{PF-Restore}\right) \approx \boldsymbol{F} \cdot \boldsymbol{S} \cdot \left[\boldsymbol{J}_{Rebin} + \beta \cdot \boldsymbol{R}_{Rebin}\right]^{-1} \cdot \boldsymbol{J}_{Rebin} \cdot \boldsymbol{S}^T. \tag{4.19}$$

Equations (4.18) and (4.19) could be used to evaluate the resolution-variance tradeoffs and bias-variance tradeoffs attainable with reconstructions using the NUOP approach. In our previous

study [94], we have demonstrated that the image quality obtained in the target-region with the NUOP approach is essentially unchanged from that obtained with uniform object-space pixelation, but the computation load could be greatly reduced, due to the decreased dimension of the object-space. This improvement is crucial for the imaging system to be a practical imaging tool for adaptive imaging study.

4.4 Monte Carlo Simulation Results

4.4.1 Introduction to a Simulated SPECT System and Phantom

To evaluate the proposed NUOP approach, a series of Monte Carlo simulations were performed. The simulated SPECT system has the same geometry as that used in the previous chapter, seen in Figure 3.4. The object-space was divided into 128×128×128 cubic voxels of 64 μm × 64 μm × 64 μm in size. In this system six detectors are arranged in a stationary hexagonal ring. Each detector has 128×128 square pixels of 350 μm × 350 μm, which is coupled to a collimation aperture with 5×5 pinholes of 200 μm in diameter. The axes of all pinholes are perpendicular to the aperture surface. The pinhole-spacing is roughly 4.2 mm. All pinholes have sharp knife-edges and an acceptance angle of 45 degrees on both sides. In the simulation the pinhole response was modeled with the analytical formula given in [97, 98]. The detector-to-aperture and aperture-to-center distances are 2.2 cm and 2 cm respectively. The depth-of-interaction effect in detectors was accounted for by treating the detector as multiple independent layers [99]. In this study, both the photon attenuation in the object and the effect of Compton scattering were ignored.

Two simulated phantoms were used in this study. The first one (resolution phantom) has a spherical volume of 8 mm in diameter placed at the center of the ring-shaped SPECT system. It has a uniform background activity concentration of A_{B1}. The central spherical volume of 3 mm diameter has a background activity concentration of A_{B2}. This volume is divided into three sections, of 1.2 mm, 0.6 mm and 1.2 mm in height, along the x-axis (the axis of the ring-SPECT system). The top section contains seven hot rods of 250 μm in diameter, superimposed on the continuous background (A_{B2}). These hot rods are parallel to the x-axis and are separated by 500 μm between adjacent rods. The bottom section has a similar configuration, except that the hot rods are replaced by cold rods of 400 μm in diameter, which are separated by 800 μm spacing. The middle section has the uniform background (A_{B2}) only. The hot rods and cold-rods have activity concentrations of A_{HR} and A_{CR} respectively. The ratios of the activity concentrations

within different regions in the simulated phantom are A_{B1}:A_{B1}:A_{HR}:A_{CR}=1:6:60:0. The entire 8 mm diameter sphere contains a total activity of 250μCi. Cross sections of the phantom for cold and hot rods are shown in Figure 4.4.

The second phantom (brain phantom) also has a spherical volume of 8 mm diameter. It consists of two hot spheres of 1 mm in diameter and two hot ellipsoidal volumes that have half-axes of 0.8 mm, 0.4 mm and 1.0 mm respectively. These features are superimposed on a uniform background and the feature-to-background ratio is 6:1. In addition, several small hot and cold spheres were inserted into the ellipsoidal volumes. The relative tracer concentrations in these small spheres are 12:1 or 1:1 in respect to the continuous background. The phantom contains a total activity of 100 μCi. The cross-sections of the phantom are shown in Figure 4.7.

In the simulation, we implemented and compared five different NUOP strategies, along with original uniform pixelation strategy. Derived using (4.13), we used the FIM-based and distance-based criteria to determine different pixelation density functions, which are shown in Figure 4.4, as long as the corresponding parameters. $u1$ and $u2$ control the degree of the relative-importance-dependency and distance-dependency respectively. Factor t of 1 would result in more aggressive reduction of pixelation density function than t of 0.5. In this work, we used an identical value s=2 for all studies.

4.4.2 Fisher Information Matrix with Non-Uniform Object-Space Pixelation

In this study, we used the resolution phantom. Since Fisher information matrix (FIM) plays an important role in determining image property, we first compared the FIMs results from the use of these PDFs. For each FIM, we computed a single column ($\boldsymbol{J}\cdot\boldsymbol{e}_j$) corresponding to the central voxel. As previously discussed, each element of this vector represents the weighted correlation between the system responses to gamma ray emissions in the given voxel and in other voxels inside the object-space [94]. The elements of $\boldsymbol{J}\cdot\boldsymbol{e}_j$ were then rearranged into 3-D format with the same order as that for positioning corresponding source voxels in spatial domain. This procedure produces a 3-D "image" of the column vector, $\boldsymbol{J}\cdot\boldsymbol{e}_j$, which is referred to as a FIM image.

In Figure 4.3, we compared 1-D cross-sections of several FIM images derived using different PDFs. Since the resultant FIM images have different overall dimensions, this comparison was limited to the FIM elements that are corresponding to those source pixels inside the spherical target-region only. The rebinning processes resulted in virtually unchanged FIM values for these voxels. The comparison of the cross-sections along different directions also highlighted the effect of the non-isotropic sampling due to the specific system geometry. Despite this similarity, the rebinning process reduces the overall number of unknowns in reconstruction and helps to improve the condition of the inverse problem. Given the same amount of imaging information carried in projection data, it will be shown later that these five PDFs could produce virtually identical images inside the ROI compared with original uniform pixelation, which reflects the FIM image similarity.

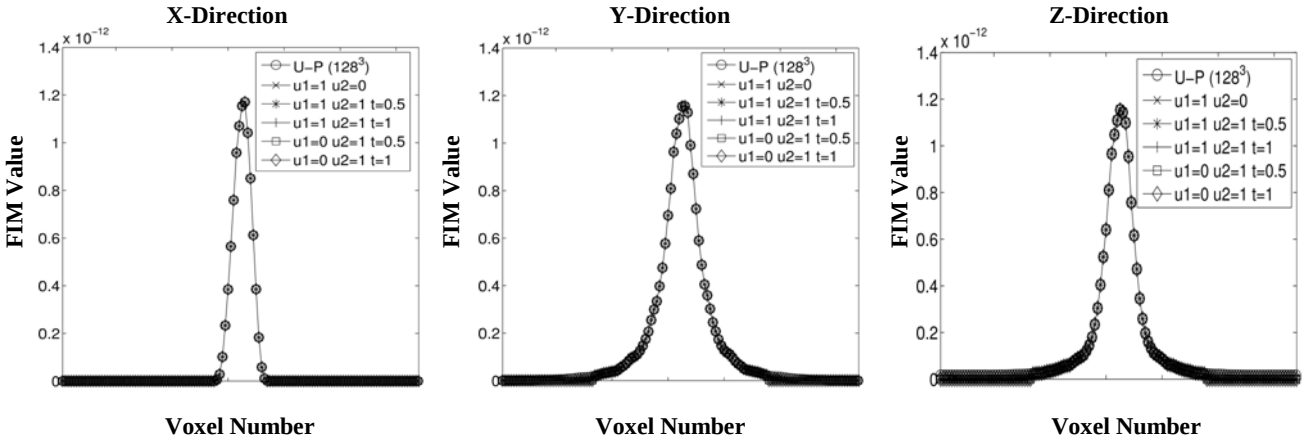

Figure 4.3 Cross-sections of several FIM images derived with different PDFs. Note that these FIMs were evaluated using the system response functions after rebinning. The NUOP process has virtually no effect on the values of the FIM elements compared.

4.4.3 Image Reconstructions with Non-Uniform Object-Space Pixelation

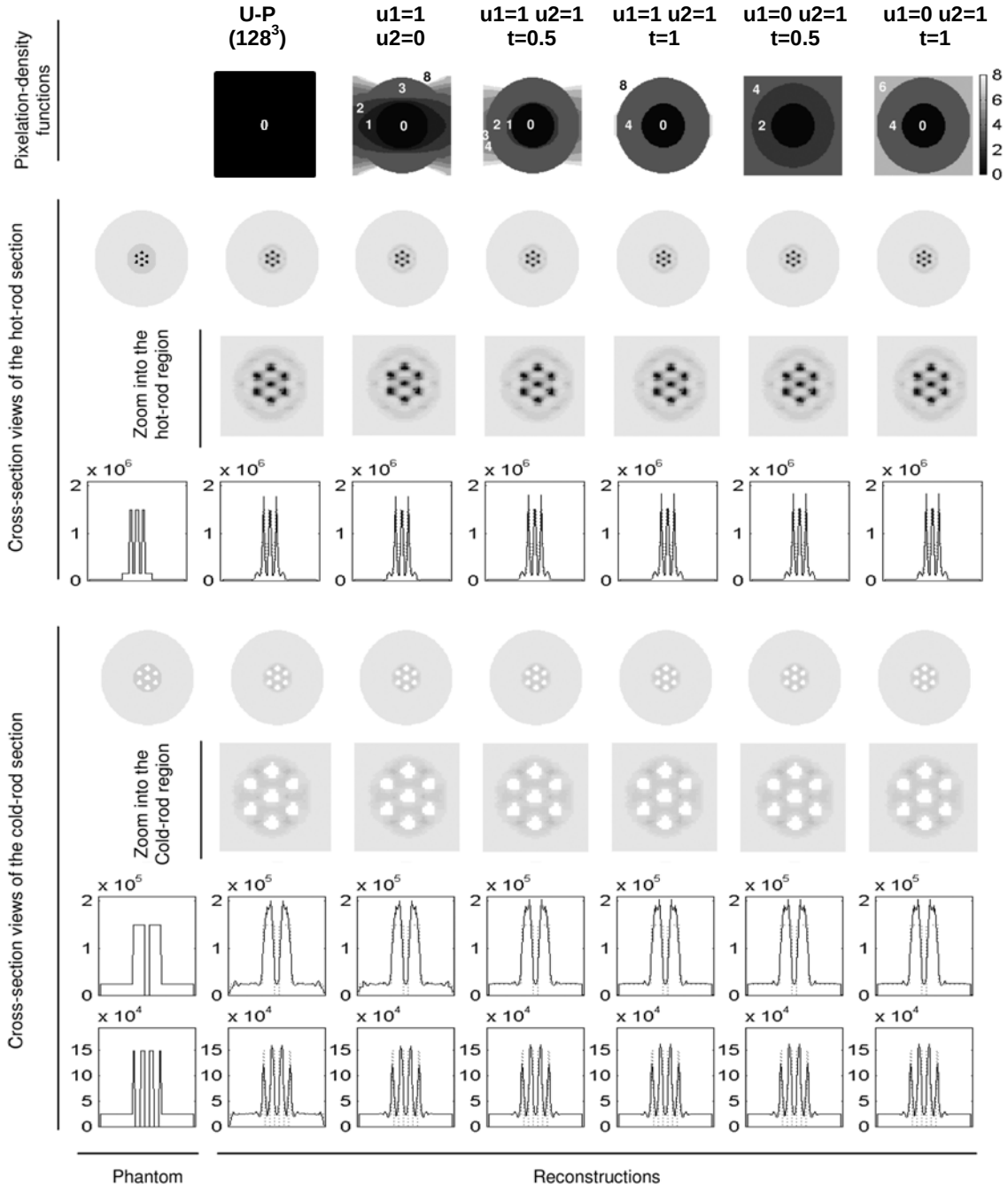

Figure 4.4 Comparing reconstructed images with different object-space pixelation schemes and noise-free projection data. The pixelation-density functions used are indicated in the top row and the parameters used for each pixelation density function are indicated. U-P: Uniform-pixelation. The 2-D slices shown are perpendicular to the common axis.

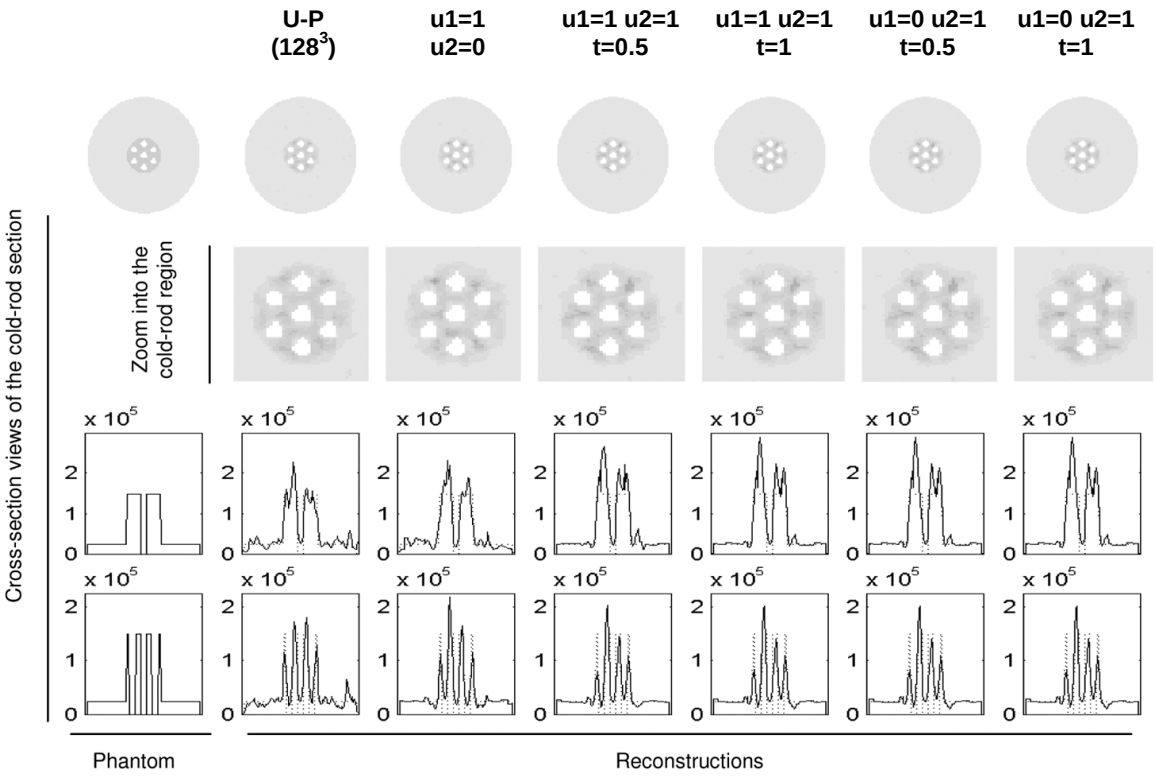

Figure 4.5 Comparing reconstructed images with different object-space pixelation schemes and noisy projection data

In this study, the central spherical region of 4 mm in diameter in the resolution phantom was defined as the target-region. Five different PDFs and uniform pixelation scheme were used and corresponding reconstructed images with noise-free projection data were compared in Figure 4.4. All images were reconstructed with 500 iterations to ensure convergence. The regularization factor β was set to 0 for all cases. The use of different rebinning strategies produced virtually identical images. A similar comparison using noisy projection data is shown in Figure 4.5, in which the different pixilation strategies resulted in comparable reconstructions. All images were reconstructed using penalized ML algorithm with 500 iterations. The regularization factor β was 10^{-14} for all reconstructions. Post-filtering operation was not used in this study.

We further evaluated the NUOP scheme with the brain phantom and randomly generated noisy projections. A rapid reconstruction, with the object divided into $32 \times 32 \times 32$ voxels, was performed to provide an overview of the object. Based on e resultant image, we defined two different target-regions (T-Rs) around the elliptical feature on the right-hand side (as shown in). Both T-Rs are elliptical and had the same orientations as the corresponding elliptical source features themselves. Their half-axes were 1.2 mm, 0.8 mm, 1.5 mm (1.5, 2 and 1.5 times the corresponding half-axes of the elliptical source feature) for the first one and 1.6 mm, 1.6 mm and

1.5 mm for the second one. Several PDFs derived for this series of studies are shown in Figure 4.6 and the reconstructed images are shown in Figure 4.7.

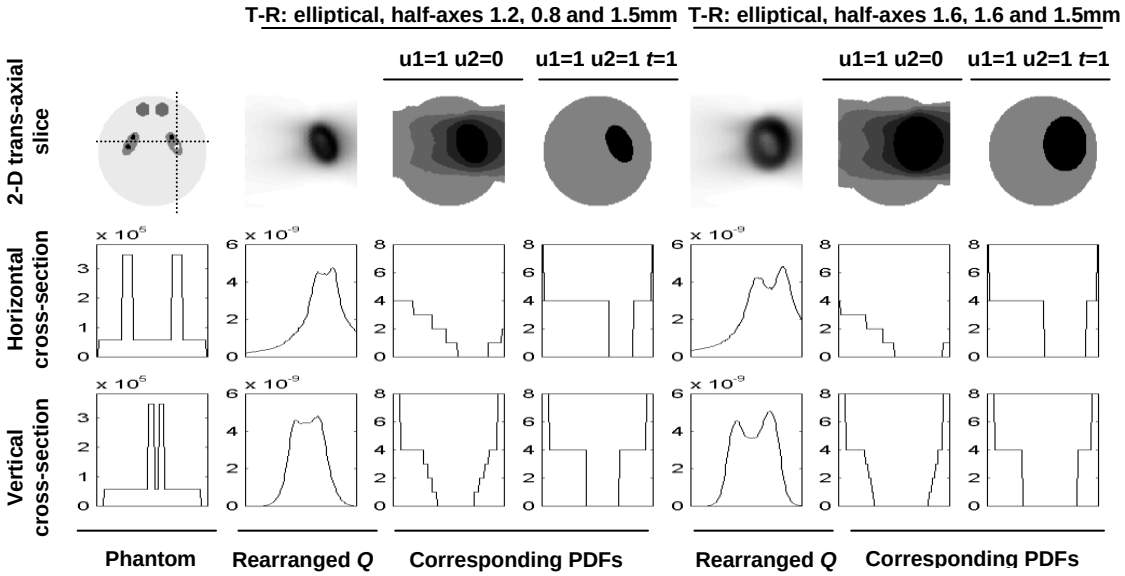

Figure 4.6 Pixelation density functions derived for differently sized target-regions in brain phantom

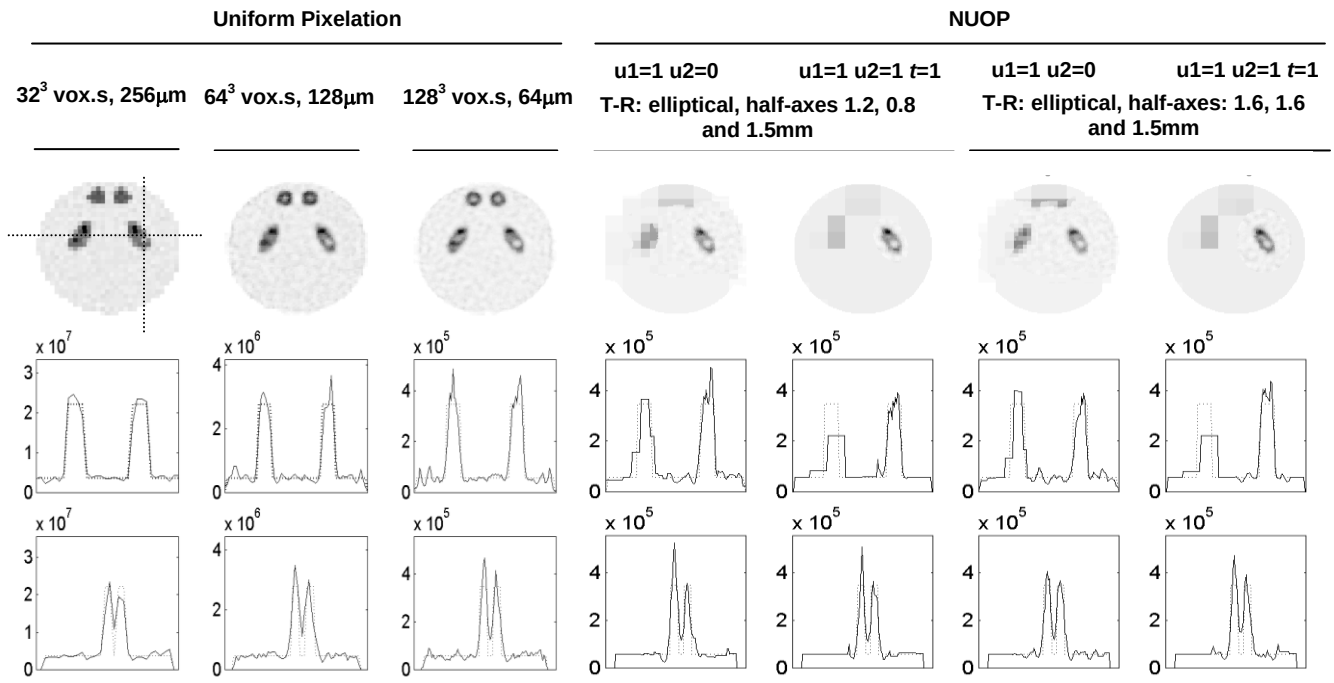

Figure 4.7 Reconstructed images of the brain phantom with different PDFs and with uniform pixelation (U-P).

4.4.4 MUCRB Studies with Non-Uniform Object-Space Pixelation

To further evaluate the impact of the NUOP process on image quality, we derived the optimum resolution-variance tradeoffs and bias-variance tradeoffs achievable with the use of different pixelation strategies, show in Figure 4.8 and Figure 4.9. This study was based on the use of the resolution phantom. The resolution-variance curves were calculated for the voxel of interest located at the center of the object. In order to derive the bias-variance curves, a spherical volume of roughly 1 mm diameter, located at the center of the object, was defined as the ROI. It contains an activity concentration twice than that of the surrounding region.

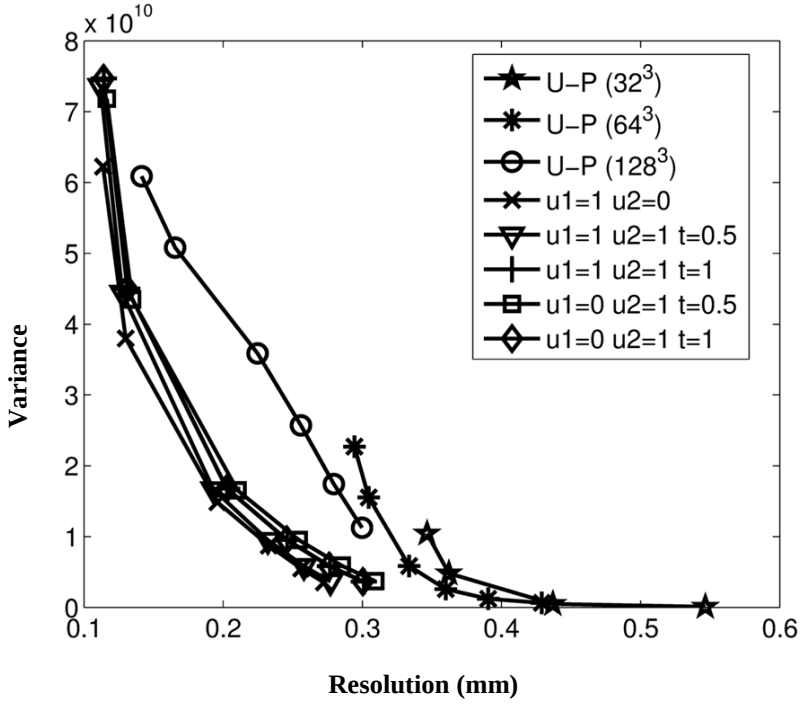

Figure 4.8 Comparison between resolution-variance curves achieved with different PDFs. Parameters used in PDFs are shown in the figure. Several resolution-variance curves achieved with uniform-pixelation (U-P) with 32^3×256 μm voxels, 64^3×128 μm voxels and 128^3×64 μm voxels are also compared.

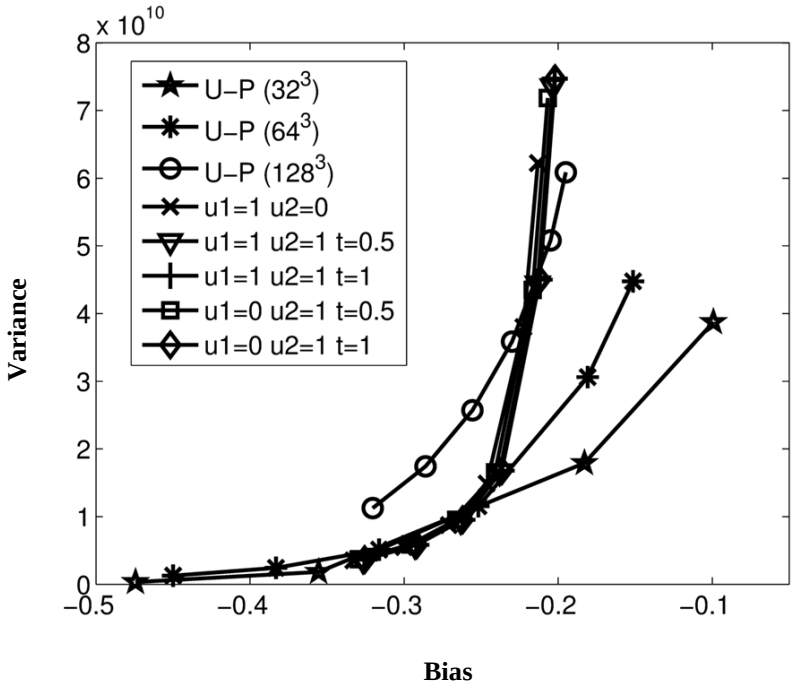

Figure 4.9 Comparison between bias-variance curves achieved with different PDFs. Several bias-variance curves achieved with uniform-pixelation (U-P) with 32^3 ×256 μm voxels, 64^3 ×128 μm voxels and 128^3 ×64 μm voxels are also compared.

In this comparison, the NUOP approach not only offered a much faster reconstruction (see Table 4.1), but also produced lower variances at similar resolutions. Furthermore, since the simulated SPECT system uses detectors with 350 μm × 350 μm pixels and pinholes of 200 μm in diameter, the practical spatial resolution of the system is expected to 200 μm – 350 μm in FWHM, which could be used to determine the desired filter matrix $\boldsymbol{F}$ and corresponding practical bias region in (3.35) and (3.36). It was shown in Figure 4.9 that the NUOP schemes also offered a clear advantage over the uniform pixelation schemes (with identical voxel size for the target-region) in terms of bias-variance tradeoffs within the practical region (mean percentage bias is less than -0.2). The definition of mean percentage bias is seen in (2.6). In this study, we also derived several resolution–variance tradeoff curves and bias–variance tradeoff curves using different uniform voxel sizes (such as 64 μm × 64 μm × 64 μm, 128 μm × 128 μm × 128 μm and 256 μm × 256 μm × 256 μm) in the object-space.

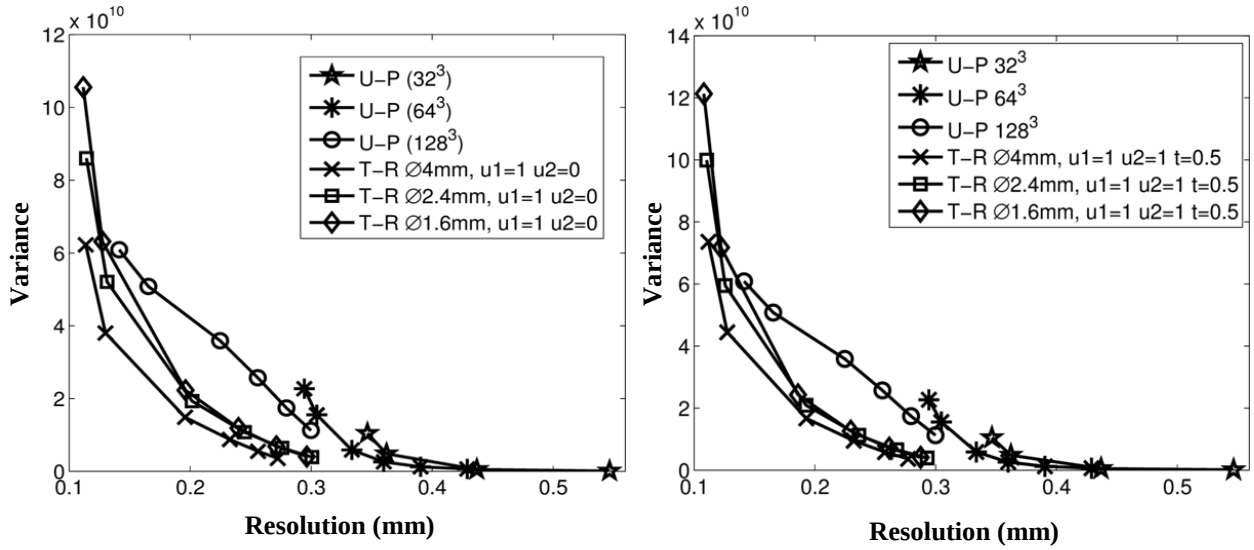

Figure 4.10 Comparisons between resolution-variance curves achieved with differently sized target-region using different types of pixelation density functions.

For implementing the NUOP approach, it is natural to question about "how large a target-region should be for a given local feature?" In this study, this question is briefly explored with a few examples as shown below. The ROI was a small spherical local region centered the voxel-of-interest. The diameter of a spherical target-region was varied from 1.6 mm, 2.4 mm to 4 mm. To make the results more representative, we used two different PDFs in this comparison. Since the definition of PDF enforces the highest pixelation density within the T-R, its dimension would have a significant influence on image reconstruction. This was confirmed by the results shown in Figure 4.10 and Figure 4.11. For the spherical object, a target-region of 4 mm diameter appeared to be the favorable choice for optimized resolution-variance tradeoffs. The use of differently sized target-regions did not have appreciable effect on bias-variance tradeoffs, and all results with NUOP are significantly better than those achieved with uniform object-pixelation within practical operating region (the mean percentage bias is less than -0.2).

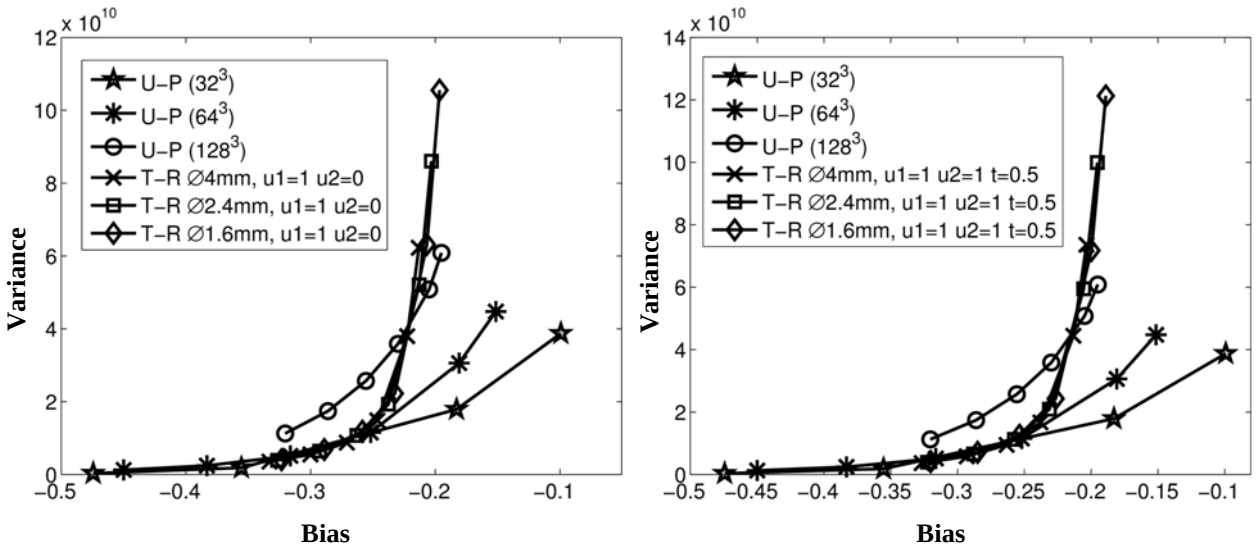

Figure 4.11 Comparison between bias-variance curves achieved with differently sized target-region using different types of pixelation density functions..

Table 4.1 the Sizes of System Response Functions and Reconstruction Times with Different Pixelation Schemes

		No. of source	Size of SRF (MB)	Reconstruction time (100 iteration)
U-P (128^3)		2097152	3950	1264 mins
T-R: Φ4mm	u1=1 u2=0	228961	509	26 mins
	u1=1 u2=1 t=0.5	132333	283	13 mins
	u1=1 u2=1 t=1	130297	274	25 mins
	u1=0 u2=1 t=0.5	142072	322	15 mins
	u1=0 u2=1 t=1	130304	275	15 mins
T-R: Φ2.4mm	u1=1 u2=0	46681	117	6 mins
	u1=1 u2=1 t=0.5	30525	74.8	4 mins
T-R: Φ1.6mm	u1=1 u2=0	15589	48.5	3 mins
	u1=1 u2=1 t=0.5	10853	35	2 mins

The sizes of system response functions for all the imaging scenarios studied are summarized in Table 4.1, along with the corresponding reconstruction times. The reconstruction times are scaled corresponding to the use of a single CPU only. It is demonstrated that the use of the NUOP approach offers greatly improved computation speed, whilst providing an excellent reconstruction for the target-region. With the rapid advances in parallel computing, images of the target region could be updated in real time during an imaging study. This feature helps to maximize the efficiency for collecting imaging information on given features, which is crucial for high resolution imaging studies.

Based on the results presented in this section, several observations and discussions are given below:

- If the target-region and the desired voxel-size in the region are fixed, the use of NUOP could offer a much reduced computation effort. Furthermore, with the same amount of imaging information carried in projection data, the use of NUOP could lead to better tradeoffs between resolution (bias) and variance. In reality, reconstruction in the non-uniformly pixelated object-space could be considered as another discrete estimation of the true source using the same projection data as that used in the uniformly pixilated object-space. In process of restoration, those voxels combined together have the same value which is equal to that of the corresponding rebinned voxel. Therefore, it is equivalent for the NUOP approach to impose constraints that all the voxels combined together are restricted to be the same. As long as rebinning strategies make sense, this method may provide a more accurate model of the true source and therefore, the image quality inside the target-region could be possibly improved. In another word, given the fixed amount of imaging information carried by projection data, reduced overall number of unknowns may help to improve the condition of the inverse problem and thus to reduce noise in reconstructed images of the target-region. However, in practice true source is unknown and rebinning strategy has to be determined based on a rough image estimate involved with noise, which may result in the inappropriate constraints and aliasing errors. Therefore the system performance may become worse, which deserves some future studies.

- The voxel size that delivers the best resolution-variance curves is not necessarily optimal for ROI quantification or other imaging tasks. To derive the tracer uptake in a ROI of 4 mm diameter, the use of 256 μm × 256 μm × 256 μm voxel size (32×32×32 voxels) produced the lowest variance at the same bias. In order to observe the ROI which could be considered as a very large voxel, coarse pixelation may still provide sufficient resolution information but greatly reduce variance and thus obtain the good balance. Clearly, the choice of voxel size should depend on the specific imaging task. The NUOP formulation developed in this work provides a computationally efficient approach for future explorations on this topic.
- In this comparison, we implement three types of NUOP strategies, based on the FIM only, FIM and distance, and distance only. All the NUOP strategies led to similar resolution-variance tradeoffs. For the SPECT systems simulated, their geometries are so designed that most of source voxels strongly correlated to the target region are physically located close to the target. Therefore, both FIM- and distance-based criteria allow most of these "important" voxels to be retained without rebinning. This leads to the comparable resolution-variance

performances shown in Figure 4.8 and Figure 4.9. The key advantage of the FIM-based approach over the distance-based one is that the former can be adapted to any object geometry with target-regions of any shape and size, whilst the latter does not offer the same degree of flexibility.

- In the comparison between the bias-variance curves for the uniform and non-uniform pixelation cases (both having the same pixel-size of 64 μm × 64 μm × 64 μm for the target-region), the uniform pixelation scheme provided lower variances at relatively small biases (for example, when the mean bias percentage is larger than -0.2 in Figure 4.9 and Figure 4.11). However, within this operating region, the resultant imaging resolutions are close to the physically permitable limit for the given system design. Therefore the corresponding reconstructions suffer from excessive noise amplification as indicated in the resolution-variance curves shown in Figure 4.8 and Figure 4.10. For the simulated SPECT systems that are based on detectors with 350 μm×350 μm pixels and pinholes of 200 μm diameter, we would expect the system resolution to be above 200 μm. If one considers only the data points that are associated with "practical" spatial resolutions (> 200μm), the use of NUOP schemes offered superior bias-resolution tradeoffs.

4.5 Summary

We proposed and evaluated a non-uniform object-space pixelation (NUOP) method for image reconstructions. Several approaches for defining the pixelation-density function were discussed and evaluated using Monte Carlo simulations. The impact of using the NUOP method was also studied based on the optimum resolution-variance tradeoffs and bias-variance tradeoffs achievable with different pixelation strategies. If the imaging task is to achieve an excellent reconstruction of a local target-region, the use of NUOP provides a greatly improved reconstruction speed. Combined with efficient parallel computers equipped with adequate memory space, practical SPEM images could be reconstructed within several minutes, rather than a few days as in our current practice.

This development benefits the task of imaging system optimization. In practice, since the object is generally unknown, choosing the optimum configuration in advance could be problematic even for experienced users. Instead, an imaging study could be started by acquiring the preliminary imaging information, which can be used to refine the imaging task and help to optimize the system configuration in the real time based on several statistical approaches that predict the performance of the system for the specific imaging task. To ensure a meaningful

decision on the “optimum” system configuration, many performance measures developed so far require the mean and covariance of reconstructed images. The evaluation of these quantities typically involves the inversion of Fisher information matrix (FIM). With the regular uniform object pixelation scheme, this procedure is normally too computation-intensive to use in the adaptive imaging method. In comparison, the non-uniform pixelation schemes developed in this section allow for rapid evaluations of the mean and covariance of reconstructed images, which could be considered as an important step towards a feasible system performance optimization.

CHAPTER 5

ADAPTIVE ANGULAR SAMPLING FOR SPECT IMAGING

One of the recent emphases in SPECT instrumentation is to push for higher spatial resolution. In recent years, many state-of-art small animal SPECT systems with excellent spatial resolution properties have been designed and constructed [33-44]. However, the biggest disadvantage of high resolution SPECT system is its relatively low collection efficiency. As a result, the long imaging time is often required for ultra-high resolution studies, which could preclude many interesting applications.

This problem could be partially alleviated by using the adaptive imaging concept proposed by Barrett et al. [47], and Freed et al. [49]. In an adaptive SPECT system, the system hardware could vary in real-time to maximize the efficiency for collecting useful imaging information regarding a given task, and therefore provide an optimum imaging performance. A reasonable adaptation rule, in reality a practical optimization method, is essential in order to realize the adaptive data acquisition.

Suppose a SPECT imaging system is characterized by a total of S system parameters, $\boldsymbol{q}=[q_1, q_2, \ldots, q_S]$. These parameters could include detector and aperture positions, detector intrinsic resolution, configurations of pinholes on each aperture and many more. The optimization task should be written in a mathematic form as:

$$\boldsymbol{q}_{optimum} == \underset{\boldsymbol{q}=[q_1,q_1,\cdots q_S]}{\arg\max}\,[\Omega(\boldsymbol{q})]. \tag{5.1}$$

The system performance index Ω is typically a scalar or a vector function of $\boldsymbol{q}$. In our study, Ω could be chosen as resolution-variance tradeoffs across multiple voxels or bias-variance tradeoffs for ROI quantification.

In this chapter, the optimization aspect is limited to imaging time distribution for a rotating SPECT system, which builds a foundation for further optimization with respect to many other parameters. In our attempts of using the SPECT imaging system for *in vivo* tracking radiolabeled cells in mouse, we often encounter a common situation – the ROI is often superimposed on or located near other strong background features. For example, SPECT imaging of pancreatic beta cells is often interfered by the presence of much stronger background from the bladder, seen in Figure 5.1. The signal from radiolabeled T cells in the brain is often very small compared to the background from other structures nearby as well. In such situations, the conventional uniform

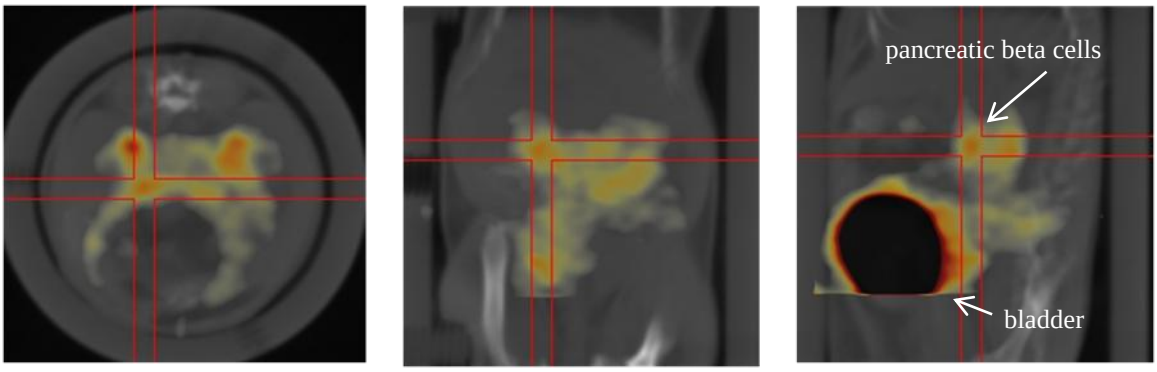

Figure 5.1 SPEM images of a small mouse (20g weight) acquired after administration of ^{99m}Tc-DTPA-Glipizide (1mCi/mouse IV injection) to evaluate beta cell loss in pancreatic disease.

angular sampling approach may not provide the most efficient way for gathering imaging information related to the ROI. Thus in this section we propose an analytical approach for performing adaptive angular sampling in SPECT imaging. This approach allows the camera to spend larger fractions of imaging time at angles those are relatively more important for acquiring useful imaging information regarding a given imaging task. With the adaptive angular sampling approach, an imaging study could start with a uniform time distribution across all possible angles. During the study, the imaging information being acquired and the input from the user (e.g., the target-region to be examined) will be used to determine the optimum time distribution based on the expected system performance measured with certain analytical performance indices [50].

The adaptive angular sampling approach requires an efficient computation method for searching through all the possible time distributions to find the optimum one in real time, which should deliver the best system performance, e.g., the minimum variance at a given resolution. To allow for a rapid optimization process, we have incorporated the vector modified uniform Cramer-Rao bound (MUCRB) approach [73] and non-uniform object-space pixelation (NUOP) approach [94] to nearly instantly evaluate the system performance. Furthermore, we have proposed a search algorithm that utilizes the gradient function of certain system performance indices, such as image variance, with respect to imaging times at individual angles. By combining the gradient-based iterative search algorithm with the rapid system performance evaluation, one can refine the angular sampling strategy adaptively during an imaging study in almost real time to achieve an improved image quality.

5.1 MUCRBs for the Adaptive Angular Sampling Approach

In this study, we chose to determine the optimum angular sampling strategy based on the resolution and (co)variance properties of reconstructed images. More specifically, it is chosen to be the one that delivers the lowest image variance at given spatial resolutions. The optimization process is implemented based on the vector modified uniform Cramer-Rao bounds (MUCRBs) [72]. In this chapter, we will modify the original vector MUCRB formulations to incorporate the proposed adaptive sampling strategy.

Suppose one uses a single-head SPECT system for imaging an object from a total of K sampling angles. The imaging time spent at k'th sampling angle is denoted by $t^{(k)}$, and thus $\boldsymbol{t}=[t^{(1)}, t^{(2)} \ \cdots \ t^{(K)}]^{\mathrm{T}}$ could represent a non-uniform imaging time distribution. Different from (2.1), imaging time $\boldsymbol{t}$ is specified in the formulas of mean projections, given by

$$\begin{aligned} &\bar{\boldsymbol{y}}^{(k)} = t^{(k)} \cdot \bar{\boldsymbol{p}}^{(k)}, \\ &\text{and} \\ &\bar{\boldsymbol{p}}^{(k)} = \boldsymbol{A}^{(k)} \cdot \boldsymbol{x} + \boldsymbol{r}^{(k)}, \ k = 1, \cdots, K. \end{aligned} \tag{5.2}$$

$\bar{\boldsymbol{p}}$ and $\boldsymbol{r}$ p̄are mean projection and mean background noise respectively within a unit imaging time. The unknown parameter $\boldsymbol{x}$ therefore represents the tracer activity in the object-space. Note that the vectors and matrices with superscript (k) are corresponding to observing the object with a single camera placed at the k'th angle only. In this study, the Fisher information matrix is defined explicitly as a function of the imaging times, with its elements given by

$$\boldsymbol{J}_{ij}(\boldsymbol{t}) = -E\left[\frac{\partial^2}{\partial x_i \partial x_j} L(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{t}) \right]. \tag{5.3}$$

The corresponding log-likelihood function can be written as:

$$\begin{aligned} L(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{t}) &= \sum_{m=1}^{M} \left(y_m \log \bar{y}_m - \bar{y}_m \right) = \sum_{k=1}^{K} \left[\sum_{m=1}^{M^{(k)}} \left(y_m^{(k)} \log t^{(k)} \bar{p}_m^{(k)} - t^{(k)} \bar{p}_m^{(k)} \right) \right] \\ &= \sum_{k=1}^{K} \left[L^{(k)}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{t}) \right], \end{aligned} \tag{5.4}$$

where

$$L^{(k)}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{t}) \equiv \sum_{m=1}^{M^{(k)}} \left(y_m^{(k)} \log t^{(k)} \bar{p}_m^{(k)} - t^{(k)} \bar{p}_m^{(k)} \right) \tag{5.5}$$

$L^{(k)}(\boldsymbol{x},\boldsymbol{y},\boldsymbol{t})$ is the log-likelihood function for the measurement acquired with the camera at the k'th angle only. Substituting (5.4) into (5.3), the FIM for non-uniform angular sampling strategy is given by:

$$\boldsymbol{J} = \sum_{\boldsymbol{k=1}}^{\boldsymbol{K}} \boldsymbol{J}^{(k)} = \sum_{k=1}^{K} t^{(k)} \boldsymbol{J}^{(k)(0)}, \tag{5.6}$$

where

$$\boldsymbol{J}^{(k)(0)} = (\boldsymbol{A}^{(k)})^{T}\, diag\left\{\frac{1}{\bar{p}^{(k)}}\right\} \cdot \boldsymbol{A}^{(k)}. \tag{5.7}$$

$\boldsymbol{J}^{(k)(0)}$ is called the elementary Fisher information matrix that has measurements acquired within a unit time spent at the k'th angle only. Substituting (5.6) into (3.31) and (3.32), the covariance of images reconstructed with post-filtered penalized maximum likelihood algorithm, given by (2.6), can be approximated as

$$\begin{aligned} Cov[\hat{\boldsymbol{x}}] \approx \boldsymbol{F} \cdot \left[\left(\sum_{k=1}^{K} t^{(k)} \boldsymbol{J}^{(k)(0)}\right) + \beta \cdot \boldsymbol{R}\right]^{-1} \cdot \left(\sum_{k=1}^{K} t^{(k)} \boldsymbol{J}^{(k)(0)}\right) \\ \cdot \left[\left(\sum_{k=1}^{K} t^{(k)} \boldsymbol{J}^{(k)(0)}\right) + \beta \cdot \boldsymbol{R}\right]^{-1} \cdot \boldsymbol{F}^{\boldsymbol{T}}. \end{aligned} \tag{5.8}$$

The variance on the j'th voxel could therefore be approximately given by,

$$\begin{aligned} Var[\hat{\boldsymbol{x}}_j] \approx \boldsymbol{e}_j^{T} \cdot \boldsymbol{F} \cdot \left[\left(\sum_{k=1}^{K} t^{(k)} \boldsymbol{J}^{(k)(0)}\right) + \beta \cdot \boldsymbol{R}\right]^{-1} \cdot \left(\sum_{k=1}^{K} t^{(k)} \boldsymbol{J}^{(k)(0)}\right) \\ \cdot \left[\left(\sum_{k=1}^{K} t^{(k)} \boldsymbol{J}^{(k)(0)}\right) + \beta \cdot \boldsymbol{R}\right]^{-1} \cdot \boldsymbol{F}^{\boldsymbol{T}} \cdot \boldsymbol{e}_j, \end{aligned} \tag{5.9}$$

where $\boldsymbol{e}_j$ is the j'th unit vector. The mean gradient matrix of the corresponding reconstruction can be approximated as

$$\boldsymbol{G}(\boldsymbol{x}) = \frac{\partial E(\hat{\boldsymbol{x}})}{\partial \boldsymbol{x}} \approx \boldsymbol{F} \cdot \left[\left(\sum_{k=1}^{K} t^{(k)} \boldsymbol{J}^{(k)(0)}\right) + \beta \cdot \boldsymbol{R}\right]^{-1} \cdot \left(\sum_{k=1}^{K} t^{(k)} \boldsymbol{J}^{(k)(0)}\right). \tag{5.10}$$

The j'th column vector of $\boldsymbol{G}(\boldsymbol{x})$ represents the local impulse response function centered at the j'th voxel in the object-space [100], obtained with the reconstruction process (2.6). Based on (5.8) and (5.10), we can compare given time distributions based on the minimum imaging variance attainable at given spatial resolutions. This method, typically referred to as the resolution-variance tradeoff, has been used extensively for optimizing the design of emission tomography systems. As previously discussed, value of β in (5.10) should be very small to satisfy the

resolution constraint (3.19). As a comprehensive resolution description, mean gradient matrix should then be very close to the desired filter matrix $\boldsymbol{F}$ and have not significant relationship with different non-uniform time distributions. Therefore we could predefine the resolution property by $\boldsymbol{F}$ at first, and then only minimize the corresponding variance with respect to imaging times.

In addition, one could also compare different imaging configurations based on their performance for ROI quantification. With the boundary of the ROI defined as (3.33), one can substitute the FIM for non-uniform time distribution into (3.35) and (3.36), and approximately derive the variance and bias associated with tracer-uptake in the ROI $tr(\hat{\boldsymbol{x}})$ as

$$Var_{ROI}\left[tr(\hat{\boldsymbol{x}})\right] \approx \boldsymbol{z}^{T} \cdot \boldsymbol{F} \cdot \left[\left(\sum_{k=1}^{K} t^{(k)} \boldsymbol{J}^{(k)(0)}\right) + \beta \cdot \boldsymbol{R}\right]^{-1} \cdot \left(\sum_{k=1}^{K} t^{(k)} \boldsymbol{J}^{(k)(0)}\right) \cdot \left[\left(\sum_{k=1}^{K} t^{(k)} \boldsymbol{J}^{(k)(0)}\right) + \beta \cdot \boldsymbol{R}\right]^{-1} \cdot \boldsymbol{F}^{T} \cdot \boldsymbol{z}\,, \tag{5.11}$$

$$Bias_{ROI}\left[tr(\hat{\boldsymbol{x}})\right] \approx \boldsymbol{z}^{T} \cdot \boldsymbol{F}^{T} \cdot \left[\left(\sum_{k=1}^{K} t^{(k)} \boldsymbol{J}^{(k)(0)}\right) + \beta \cdot \boldsymbol{R}\right]^{-1} \cdot \left(\sum_{k=1}^{K} t^{(k)} \boldsymbol{J}^{(k)(0)}\right) \cdot \boldsymbol{x} - \boldsymbol{z}^{T} \cdot \boldsymbol{x}. \tag{5.12}$$

In principle, one could apply (5.8)-(5.12) to derive the resolution-variance and bias-variance tradeoffs and use these quantities as the basis to find an optimum time distribution. However, even with these close-form solutions, the computation load for searching through the entire continuous imaging time domain could be too heavy for regular PC-based computation environment. To allow near real-time optimization of the imaging time distribution, we will introduce a gradient-based search algorithm for finding the optimum time distribution.

5.2 A Gradient-Based Approach for Optimizing the Angular Sampling Strategy

In order to develop a systematic approach for minimizing the image variance corresponding to the distribution of a finite total imaging time in real-time, we first formulate a solution for the partial derivatives of the variance with respect to individual imaging times. This would help us to rank the relative-importance of each sampling angle for reducing image variance.

The partial derivative of image variance on reconstructed j'th voxel, with respect to $t^{(k)}$, is given by

$$\frac{\partial}{\partial t^{(k)}} Var\left[\hat{x}_j\left(t^{(1)}\cdots, t^{(k)}, \cdots t^{(K)}\right)\right]$$
$$= \lim_{\delta\to 0} \frac{Var\left[\hat{x}_j\left(t^{(1)}\cdots, t^{(k)}+\delta, \cdots t^{(K)}\right)\right] - Var\left[\hat{x}_j\left(t^{(1)}\cdots, t^{(k)}, \cdots t^{(K)}\right)\right]}{\delta}. \tag{5.13}$$

To evaluate the partial derivatives, we first introduce a small disturbance δ to $t^{(k)}$ and the corresponding variance becomes

$$Var\left[\hat{x}_j\left(t^{(1)}, t^{(2)}\cdots, t^{(k)}+\delta, \cdots t^{(K)}\right)\right]$$
$$\approx \boldsymbol{e}_j^T \cdot \boldsymbol{F} \cdot \left[\boldsymbol{J} + \beta\cdot\boldsymbol{R} + \delta\cdot\boldsymbol{J}^{(k)(0)}\right]^{-1} \cdot \left(\boldsymbol{J} + \delta\cdot\boldsymbol{J}^{(k)(0)}\right) \cdot \left[\boldsymbol{J} + \beta\cdot\boldsymbol{R} + \delta\cdot\boldsymbol{J}^{(k)(0)}\right]^{-1} \cdot \boldsymbol{F}^{\boldsymbol{T}} \cdot \boldsymbol{e}_{\boldsymbol{j}}. \tag{5.14}$$

For small δ, one could use the matrix inversion lemma [101], which gives

$$\left[\boldsymbol{J} + \beta\cdot\boldsymbol{R} + \delta\cdot\boldsymbol{J}^{(k)(0)}\right]^{-1} \approx \left[\boldsymbol{J} + \beta\cdot\boldsymbol{R}\right]^{-1} - \delta\cdot\left[\boldsymbol{J} + \beta\cdot\boldsymbol{R}\right]^{-1} \cdot \boldsymbol{J}^{(k)(0)} \cdot \left[\boldsymbol{J} + \beta\cdot\boldsymbol{R}\right]^{-1}. \tag{5.15}$$

Substituting (5.15) into (5.13), and letting δ approach zero, the partial derivatives become

$$\frac{\partial}{\partial t^{(k)}} Var\left[\hat{\boldsymbol{x}}_j\left(t^{(1)}\cdots, t^{(k)}+\delta, \cdots t^{(K)}\right)\right]$$
$$= \boldsymbol{e}_j^T \cdot \boldsymbol{F} \cdot \begin{Bmatrix} \left[\boldsymbol{J} + \beta\cdot\boldsymbol{R}\right]^{-1} \cdot \boldsymbol{J}^{(k)(0)} \cdot \left[\boldsymbol{J} + \beta\cdot\boldsymbol{R}\right]^{-1} \\ -2\cdot\left[\boldsymbol{J} + \beta\cdot\boldsymbol{R}\right]^{-1} \cdot \boldsymbol{J}^{(k)(0)} \cdot \left[\boldsymbol{J} + \beta\cdot\boldsymbol{R}\right]^{-1} \cdot \boldsymbol{J} \cdot \left[\boldsymbol{J} + \beta\cdot\boldsymbol{R}\right]^{-1} \end{Bmatrix} \cdot \boldsymbol{F}^{\boldsymbol{T}} \cdot \boldsymbol{e}_{\boldsymbol{j}}. \tag{5.16}$$

With (5.16), one could derive the gradient of imaging variance on the *j*'th voxel with respect to the time distribution as

$$\nabla_{t=[t^{(1)}, t^{(2)}, \cdots, t^{(K)}]} \left\{Var[\hat{x}_j]\right\} = \left\{ \partial Var[\hat{x}_j] \Big/ \partial t^{(1)}, \partial Var[\hat{x}_j] \Big/ \partial t^{(2)}, \cdots, \partial Var[\hat{x}_j] \Big/ \partial t^{(K)} \right\}. \tag{5.17}$$

One could also use a similar approach to optimize the time distribution based on other performance measures, such as the tradeoffs between the variance and bias for quantifying the total uptake in a ROI. In this case, the partial derivative of the variance (on the derived ROI value) with respect to imaging time $t^{(k)}$ is given by:

$$\frac{\partial}{\partial t^{(k)}} Var_{ROI}\left[tr\left(\hat{\boldsymbol{x}}\right)\right]$$
$$= \boldsymbol{z}^T \cdot \boldsymbol{F} \cdot \begin{Bmatrix} \left[\boldsymbol{J} + \beta\cdot\boldsymbol{R}\right]^{-1} \cdot \boldsymbol{J}^{(k)(0)} \cdot \left[\boldsymbol{J} + \beta\cdot\boldsymbol{R}\right]^{-1} \\ -2\cdot\left[\boldsymbol{J} + \beta\cdot\boldsymbol{R}\right]^{-1} \cdot \boldsymbol{J}^{(k)(0)} \cdot \left[\boldsymbol{J} + \beta\cdot\boldsymbol{R}\right]^{-1} \cdot \boldsymbol{J} \cdot \left[\boldsymbol{J} + \beta\cdot\boldsymbol{R}\right]^{-1} \end{Bmatrix} \cdot \boldsymbol{F}^{\boldsymbol{T}} \cdot \boldsymbol{z}. \tag{5.18}$$

where $\boldsymbol{z}$ is the ROI indicator defined in (3.34).

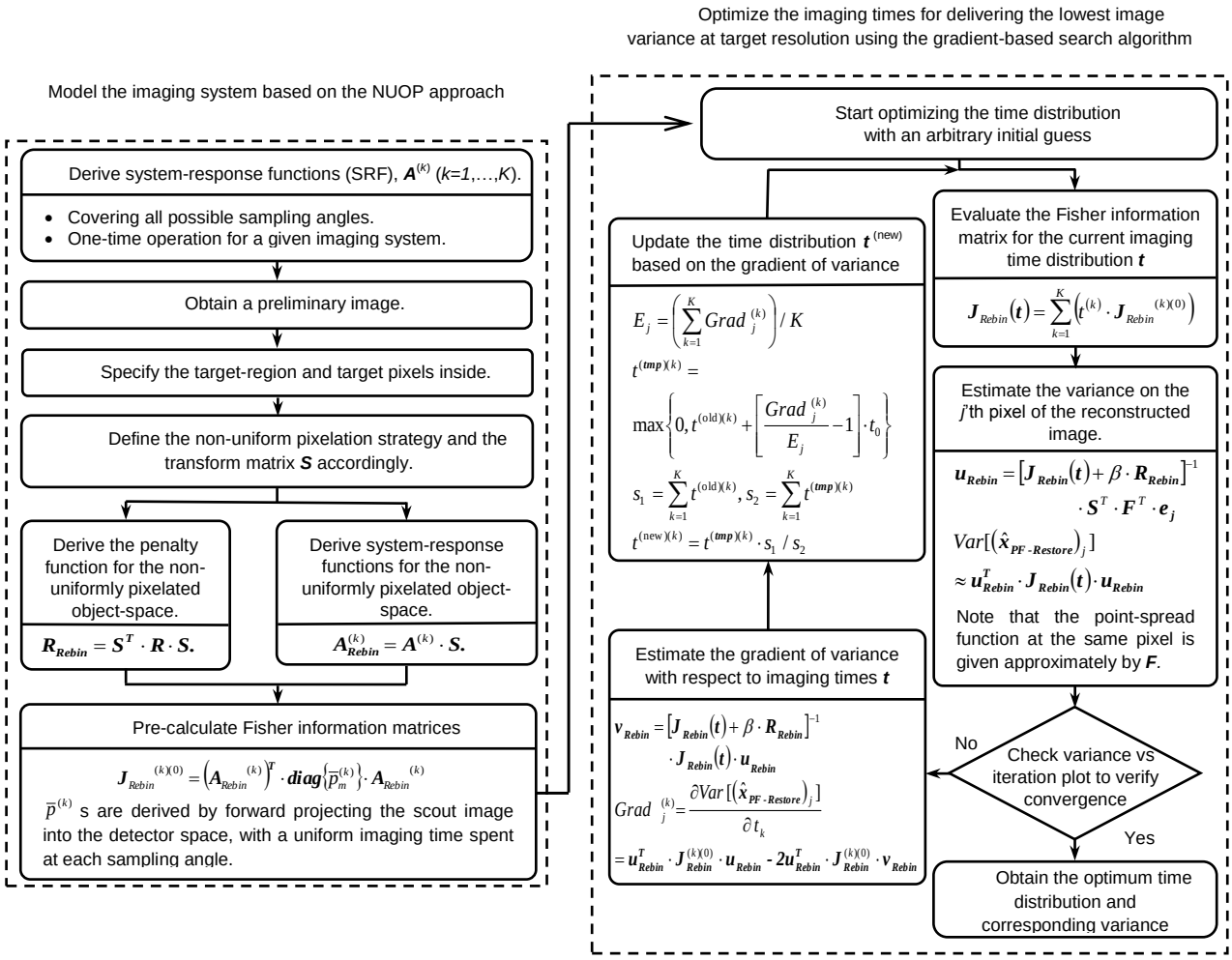

Figure 5.2 The combination of the NUOP approach and the gradient-based search algorithm for finding the optimum angular sampling strategy

To find the optimum distribution of sampling times that delivers the minimum image variance, we have proposed an iterative algorithm that uses the gradient function. At the beginning of each iteration, we start to calculate the overall Fisher information matrix and variance for the current imaging time and verify whether the variance is converged. If not, the partial derivatives of the image variance with respect to $t^{(k)}$ s are evaluated using (5.16), and then an average value across all partial derivatives is derived. To proceed to the next iteration, the imaging time for each sampling angle will be either increased or decreased, depending on the difference between the corresponding partial derivative value and the average value. This process is kept repeated iteratively, until the image variance is stabilized around its minimum value. This algorithm is sketched in Figure 5.2.

It is worth noting that the (co)variance given by (5.8) and (5.9) are not guaranteed to be convex functions of the time distribution vector $\boldsymbol{t}$. So it is possible that the proposed algorithm

converges to local minima. In principle, there could be certain object geometries possessing certain symmetry that leads to multiple sampling strategies yielding identical minimum imaging variance values.

The partial derivatives given by (5.16) can be evaluated using the following recipe. One could use conjugate gradient method [102] to compute

$$\boldsymbol{u} = \left[\boldsymbol{J} + \beta \cdot \boldsymbol{R}\right]^{-1} \cdot \boldsymbol{F}^{\boldsymbol{T}} \cdot \boldsymbol{e}_j, \tag{5.19}$$

and then

$$\begin{aligned} \boldsymbol{v} &= \left[\boldsymbol{J} + \beta \cdot \boldsymbol{R}\right]^{-1} \cdot \boldsymbol{J} \cdot \boldsymbol{u} \\ &= \left[\boldsymbol{J} + \beta \cdot \boldsymbol{R}\right]^{-1} \cdot \boldsymbol{J} \cdot \left[\boldsymbol{J} + \beta \cdot \boldsymbol{R}\right]^{-1} \cdot \boldsymbol{F}^{\boldsymbol{T}} \cdot \boldsymbol{e}_j. \end{aligned} \tag{5.20}$$

The image variance on j'th voxel is therefore

$$Var\left[\hat{x}_j\left(t^{(1)}, t^{(2)} \cdots, t^{(k)}, \cdots t^{(K)}\right)\right] \approx \boldsymbol{u}^{\boldsymbol{T}} \cdot \boldsymbol{J} \cdot \boldsymbol{u}, \tag{5.21}$$

and the partial derivative of the variance with respect to imaging time $t^{(k)}$ is given by

$$\frac{\partial}{\partial t^{(k)}} Var\left[\hat{\boldsymbol{x}}_j\left(t^{(1)} \cdots, t^{(k)}, \cdots t^{(K)}\right)\right] = \boldsymbol{u}^T \cdot \boldsymbol{J}^{(k)(0)} \cdot \boldsymbol{u} - 2 \cdot \boldsymbol{u}^T \cdot \boldsymbol{J}^{(k)(0)} \cdot \boldsymbol{v} \tag{5.22}$$

When computing imaging variance and its gradient using (5.19)–(5.22), the most computation-intensive steps are the evaluations of vectors $\boldsymbol{u}$ and $\boldsymbol{v}$, each requiring the inversion of matrix $(\boldsymbol{J}+\beta\cdot\boldsymbol{R})$. With the geometry used in the simulation later, for example, the total number of non-zero elements in is at the order of 10^9. Even with the use of sparse matrix and 8 CPUs running in parallel, it took about 3 hours for an inversion of the matrix $(\boldsymbol{J}+\beta\cdot\boldsymbol{R})$. In general, the amount of computation required for evaluating vectors $\boldsymbol{u}$ or $\boldsymbol{v}$ is similar to that for a single image reconstruction. Therefore, the total computation load for evaluating the gradient is about twice that of image reconstruction. Assuming a total of 50 iterations are needed for the gradient-based iterative approach (shown in Fig. 1) for finding the optimum time distribution, the total amount of computation will be close to that for 100 image reconstructions. This highlights the challenge for performing such optimization steps in real-time. To alleviate this problem, we have adopted the non-uniform object-space pixelation (NUOP) approach to the gradient-based iterative search algorithm with greatly reduced number of unknowns. The combined approaches offer a practical computation tool for nearly instantly performing adaptive angular sampling. The computational steps in the adaptive angular sampling approach are summarized in Figure 5.2.

5.3 Monte Carlo Simulation Results

5.3.1 Introduction to a Simulated SPECT System and Phantom

Monte Carlo studies have been performed to evaluate the adaptive angular sampling approach using a single-head SPECT system. This single camera is rotated around the object with a constant radius and *N* equally-spaced sampling angles during data acquisition, seen in Figure 5.3. In the first study for validating the adaptive angular sampling, there are totally 32 sampling angles (*N*=32) around the object during data acquisition. The gamma ray detector has 64×64 square pixels of 700 μm × 700 μm in size. It is coupled to a collimator with a single pinhole of 300 μm diameter. In the second study for further evaluating and visualizing the benefit of the adaptive angular sampling, higher spatial resolution is preferred. The camera could sample the object with 64 equally-spaced sampling angles (*N*=64). The detector is divided into 128×128 square pixels of 350 μm × 350 μm in size, and is coupled to a collimator with a single pinhole of 200 μm in diameter. In both studies, the detector-to-pinhole and pinhole-to-center distances are 45mm and 15mm respectively. The attenuation of gamma rays in collimator and the depth-of-interaction effect in detector were modeled using Monte Carlo package described in [31].

The object-space was divided into 64×64×64 voxels, each being 200 μm × 200 μm × 200 μm in size. The active volume of the object is a sphere of around 12mm in diameter, having a uniform background concentration. In the first study, we have modeled a total of 12 object configurations, each having a spherical region-of-interest (ROI) of 1.6 mm diameter and an extra ellipsoidal background region (half-axes: 1.5 mm, 1.5 mm and 3 mm) with elevated tracer updates. The ROI, the ellipsoidal background feature, and the continuous background have activity concentrations of C_1, C_2, and C_B respectively. Detailed dimensions and relative tracer uptakes of the phantoms are shown in Table 5.1 and Figure 5.4. The entire phantom contains a total activity of 500 μCi and the total observation time is 64 minutes for all MC studies. Intuitively, gamma ray attenuation in the object could play an important role in defining the optimum angular sampling strategy, especially for imaging human-sized objects with tracers labeled with Tc-99m. To highlight this effect, we have also tested three linear attenuation coefficients, 0 mm^{-1}, 0.0154 mm^{-1} (typical value for 140 keV gamma rays in water) and 0.077 mm^{-1}, in the phantoms. Please also see Table 5.1 for details. These phantoms are denoted as phantom1.

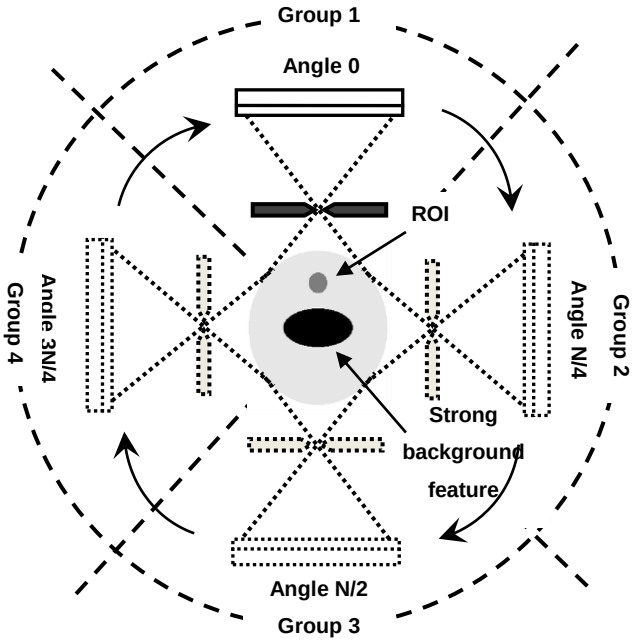

Figure 5.3 Simulated single-head SPECT system with 1-ph aperture

In the second study, the phantom configuration is slightly modified. The ROI contains three hot cylinders parallel to the rotation axis, each being 0.6 mm in diameter and 4 mm in length (instead of having a uniform and above-the-background tracer concentration). The spacing between the small spheres is 0.5 mm, seen in Figure 5.8. The off-centered strong background feature has the same size of that in the first study. The ratio of activity concentrations contained in the ROI, the ellipsoidal background feature, and the continuous background is 1:5:25. The total activity the entire phantom contains is increased to 4 mCi and the total observation time is 128 minutes. In this study, gamma ray attenuation in the object should be considered and linear attenuation coefficient in the phantom is 0.077mm^{-1}. This phantom is denoted as phantom2.

Table 5.1 Different Phantom Configurations Used to Evaluate the Proposed Computation Approach

	Phantoms with the ROI at the Center						Phantoms with the ROI off the Center					
Phantoms	P_{C1}	P_{C2}	P_{C3}	P_{C4}	P_{C5}	P_{C6}	P_{O1}	P_{O2}	P_{O3}	P_{O4}	P_{O5}	P_{O6}
Tracer Concentrations: C_B:C_1:C_2	1 : 5 : 10			1 : 5 : 25			1 : 5 : 10			1 : 5 : 25		
Linear Attenu. Coef. (mm^{-1})	0	0.0154	0.077	0	0.0154	0.077	0	0.0154	0.077	0	0.0154	0.077

As discussed before, we have used a non-uniform object-space pixelation approach to reduce the computation load. In this approach, the FIM-based criterion was used to determine the pixelation density function. The ROI and regions important to the reconstructed ROI were represented by the finest voxels and other regions were represented by larger voxels (Figure 5.5). As a result, the number of unknowns could be reduced from 262144 to around 4000. This allowed us to find the optimum angular sampling strategy for imaging the ROIs within a few minutes using a single PC.

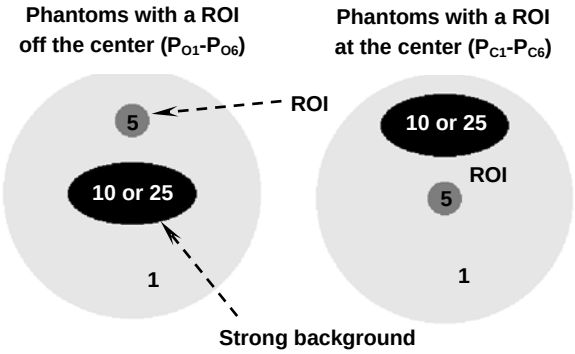

Figure 5.4 Schematics of two spherical phantoms used in the Monte Carlo simulations. Each phantom contains a small spherical region-of-interest (ROI) of 1.6 mm diameter and a strong elliptical feature superimposed on a continuous background. The location of the ROI, the relative intensities of different features and the attenuation coefficient of the media in the phantom are varied to produce a total of 12 different phantom configurations as shown in Table 4.1. The numbers shown in different regions of the phantoms are the relative tracer uptake values.

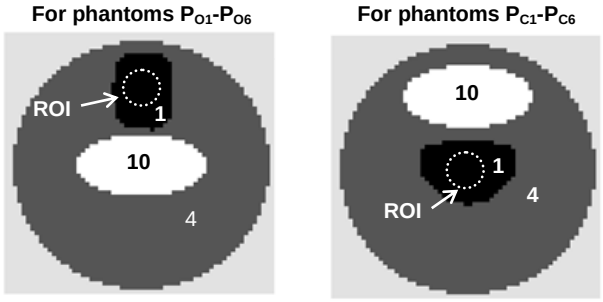

Figure 5.5 The rebinning strategy used with the phantoms. The numbers shown in the graphs are the pixelation density function. The number "1" symbolizes that the pixels in the corresponding regions (marked in black) are not rebinned. Within the background region marked with 4, 4 × 4 × 4 adjacent pixels will be rebinned into a larger pixel. In addition, pixels in the strong background region marked with number "10" will be rebinned into a single voxel. In this example, the use of the NUOP approach reduces the total number of unknowns from 262144 to around 4000.

5.3.2 Validation of the Adaptive Angular Sampling Approach

To verify the proposed gradient-based search algorithm (Figure 5.2), we have compared the optimum time distributions obtained with the proposed algorithm against the ones obtained with exhaustive searches through all possible combinations of imaging times. In this study, phantom1 with Po4 configuration was used, detailed in Figure 5.4 and Table 5.1.

To make the exhaustive search process computationally practical, we have to combine the 32 sampling angles into four groups. The first group includes angles 28-31 and 0-3, the second group includes angles 4-11 and so on (Figure 5.3). Within each group, the imaging times at all angles are kept the same. Equivalently the number of effective sampling angles is shrunk to be 4. The total imaging time for these four groups in this study is 480 seconds. To further reduce the number of possible angular sampling schemes, the imaging time for each group can only be varied as integral multiple of certain finite time intervals, such as 30s or 10s. A smaller interval would further improve the accuracy. But the greatly increased number of possible sampling configurations makes this impractical. By contrast, the proposed gradient-based search algorithm allows imaging time at each group (of eight angles) to vary arbitrarily, except that the sampling times cannot be negative. The corresponding partial derivative of variance for each group could be calculated as the sum of partial derivatives with respect to imaging times at sampling angles in the related group. The optimum time distributions derived with the exhaustive search and the gradient-based algorithm are shown in Table 5.2. Both approaches generate similar results and the degree of agreement improves with smaller time interval allowed for the exhaustive search process. β in (4.4) is set to be 1×10^{-9} and the FHWM of the corresponding 3-D Gaussian filter function used in (5.9) and (5.10) is chosen to be 0.8 mm.

Table 5.2 Comparison of Optimal Variance and Corresponding Time Distribution by Two Approaches

	Observation Time (s) Spent at				Minimum Variance[1]
	Group1	Group2	Group3	Group4	
Exhaustive Search (30s step)	270	120	30	60	376.2026
Exhaustive Search (10s step)	280	120	20	60	375.4184
Gradient-Based Approach	283.96	119.05	19.082	57.91	375.35

1. The variance with uniform time distribution is 466.0874.

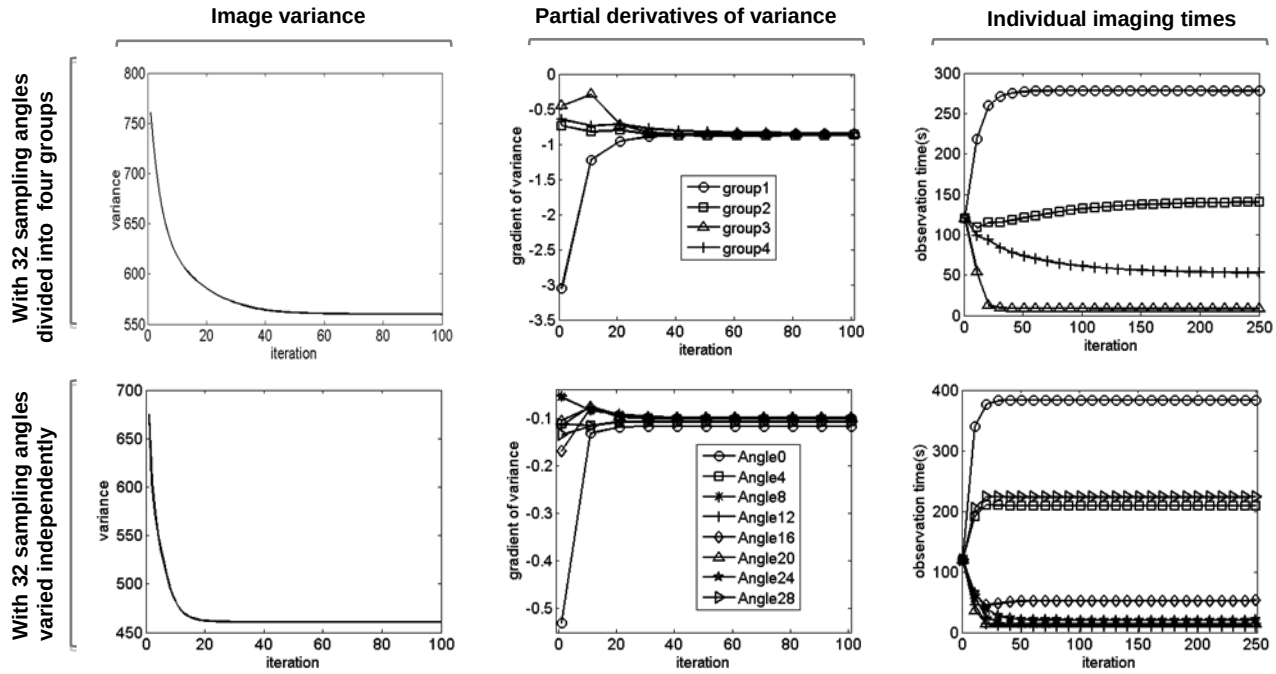

Figure 5.6 The convergence behavior of the proposed gradient-based search algorithm for finding the optimum angular sampling strategy. Top row: The image variance, its partial derivatives and the resultant imaging times at individual imaging angles (or groups of angles) as functions of iteration number. $\beta=1\times10^{-8}$. Bottom row: Similar results obtained by allow imaging times at all the 32 angles to vary independently. $\beta=1\times10^{-9}$. The FHWM of the corresponding Gaussian filter is 0.8 mm for both cases.

In this simple example, we have shown in Table 5.2 that both the exhaustive search and the gradient-based approach have led to highly non-uniform time distribution across the four (groups of) angles, which did lead to a reduced image variance by 19.5% compared with the uniform imaging time distribution. In Figure 5.6, we plot the image variance, the partial derivatives of the variance given by (5.16), and the resultant time distribution as functions of iteration number. The image variance converged to its minimum value within the first 50 iterations, but the optimum imaging times at certain angles kept changing till well beyond 150 iterations. Partial derivatives of variance with respect to imaging times approached at equal. For comparison, we have also included the results obtained by allowing all the 32 imaging times to vary independently. In this study, the non-uniform angular sampling led to a greater reduction of image variance (by ~31.8%), seen in Table 5.3, due to greater degree of time flexibility. Combination of NUOP approach and gradient-based search algorithm provided an efficient computation way to quickly find the optimum time (less than 1 minute per iteration in Figure 5.6).

5.3.3 Optimum Time Distributions for ROI Studies

We have evaluated the use of the proposed algorithm to find the optimum angular sampling strategy, especially for imaging relatively weak ROIs inside objects containing hot background features. In this effort, we studied a total of 12 phantom configurations, as detailed in Table 5.1. The imaging times at all possible sampling angles were allowed to vary independently. To ensure the variance to converge to their minimum values, we allowed 250 iterations for the optimization process. Figure 5.7 summarizes the optimum time distribution obtained using the gradient-based search algorithms for all 12 phantom configurations. In

Table 5.3, the imaging variance values obtained with the non-uniform sampling strategies are compared with those derived using uniform sampling across all 32 sampling angles.

Table 5.3 Imaging Variance Obtained with Non-Uniform Angular Sampling Strategies

Phantom	Variance (uniform time)	Variance (non-uniform time)	Relative Reduction
P_{O1}	433.37	319.05	0.264
P_{O2}	457.20	321.08	0.298
P_{O3}	549.16	334.78	0.390
P_{O4}	675.16	460.59	0.318
P_{O5}	709.29	461.19	0.350
P_{O6}	827.97	469.84	0.433
P_{C1}	504.38	481.94	0.044
P_{C2}	549.96	516.73	0.060
P_{C3}	782.95	697.03	0.110
P_{C4}	769.26	696.70	0.094
P_{C5}	835.18	739.14	0.115
P_{C6}	1154.36	953.14	0.174

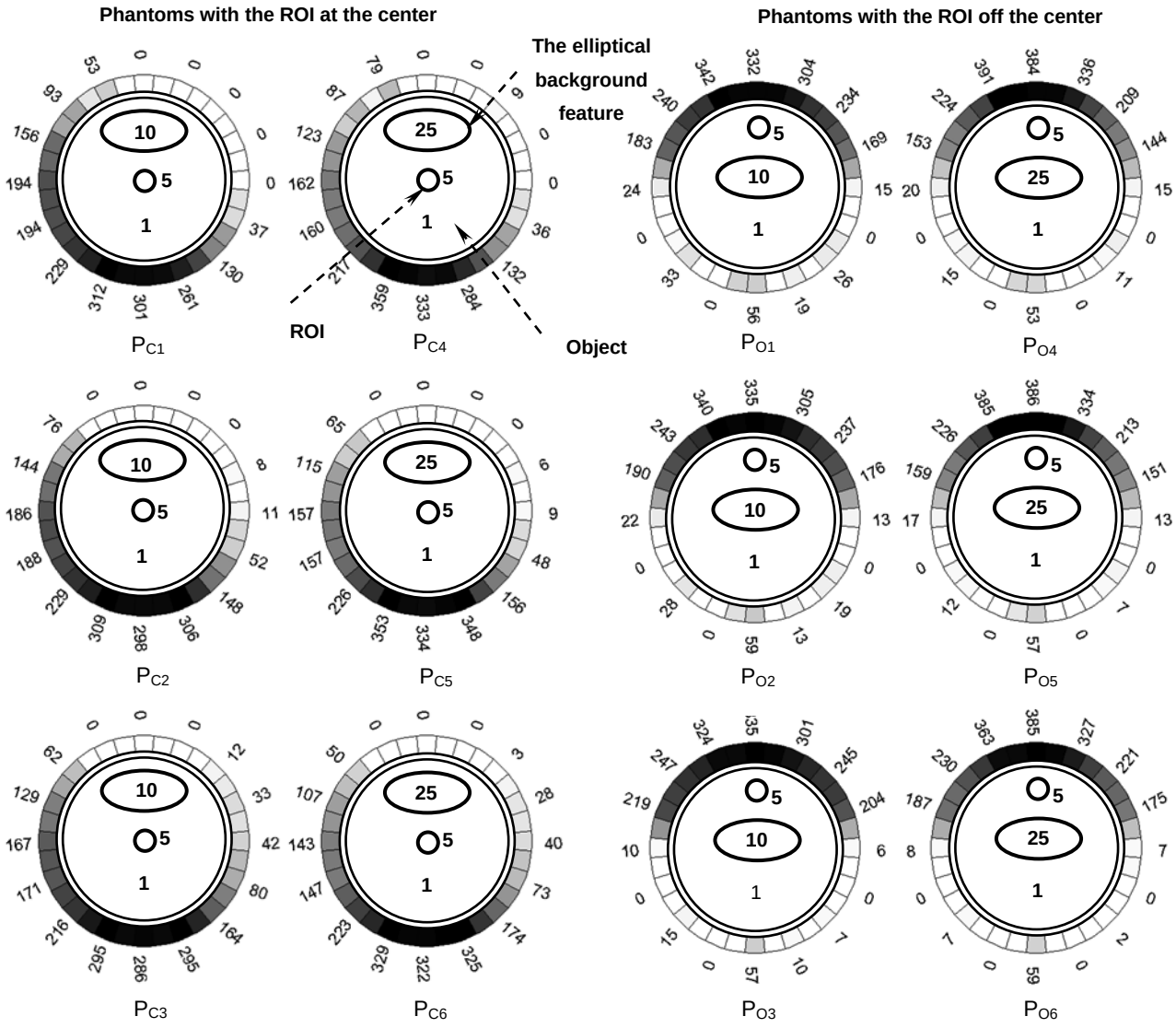

Figure 5.7 The optimum time distributions across the 32 sampling angles, obtained using the gradient-based search algorithm with 300 iterations. The color-bar and the numbers (with the unit of second) surrounding the phantom represent the optimum sampling times at individual angles. β is 1×10^{-8} for all the cases. The FHWM of corresponding Gaussian filter function is 0.8 mm.

In general, the adaptive angular sampling approaches have led to an appreciable reduction in image variance (by factors of 10-40%), over the uniform sampling approaches, particularly for imaging ROIs located near the boundary of the object (PO1-PO6). In such cases, imaging angles, corresponding to the camera placed close to the ROI, are the most effective ones for collecting imaging information. The resultant noise reduction is greater for phantoms with increased gamma ray attenuation. This effect can be seen by comparing the noise reduction factors within different groups (PO1-PO3), (PO4-PO6), (PC1-PC3), and (PC4-PC6). Furthermore, the noise reduction must be more obvious for stronger background feature, which could be seen in the comparison of groups (PO1&PO4), (PO2&PO5), (PO3&PO6), (PC1&PC4), (PC2&PC5) and (PC3&PC6).

In order to further study the adaptive angular sampling approach, phantom2 was sampled with 64 sampling angles in the study. By using the computation procedures illustrated in Figure 5.2, we have identified the optimum angular sampling strategy as shown in Figure 5.8. Compared to the uniform sampling approach, the non-uniform optimum time distribution led to a reduction in image variance by 42%. To visualize the performance benefit of using the adaptive sampling approach, we have compared two noisy images of the phantom shown in Figure 5.9 with both uniform and non-uniform angular sampling strategies. Both images had the same variance at the center of the ROI, calculated from 100 realizations of noisy reconstructions and controlled by the FWHM of Gaussian filter in the post-filtered penalized maximum-likelihood reconstruction process (2.6). Clearly, the adaptive angular sampling provides a better spatial resolution in the ROI while providing similar noise levels, which can be visualized in both 1-D and 2-D cross sections of the reconstructed images. It is shown that the performance of a SPECT system can be improved with non-uniform allocations of imaging time across all possible sampling angles. Note that both images are reconstructed in the uniform object-space. This comparison indicates that the NUOP approach could give accurate prediction for image quality.

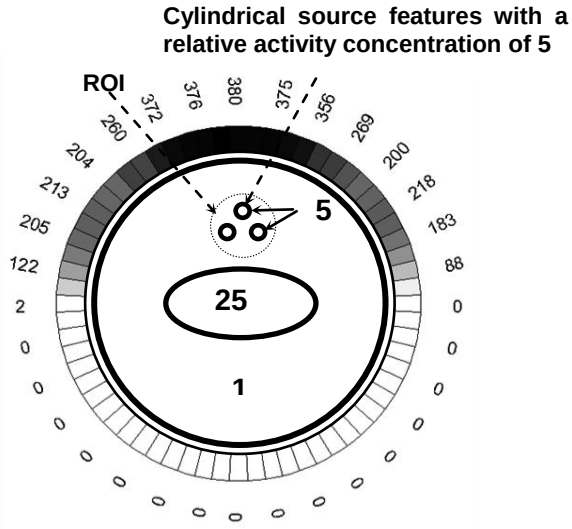

Figure 5.8 The optimum time distributions across the 64 sampling angles for imaging the three hot tubes. The numbers marked inside the object are the relative activity concentrations of different features and the background. The numbers around the color bar are imaging times given in seconds.

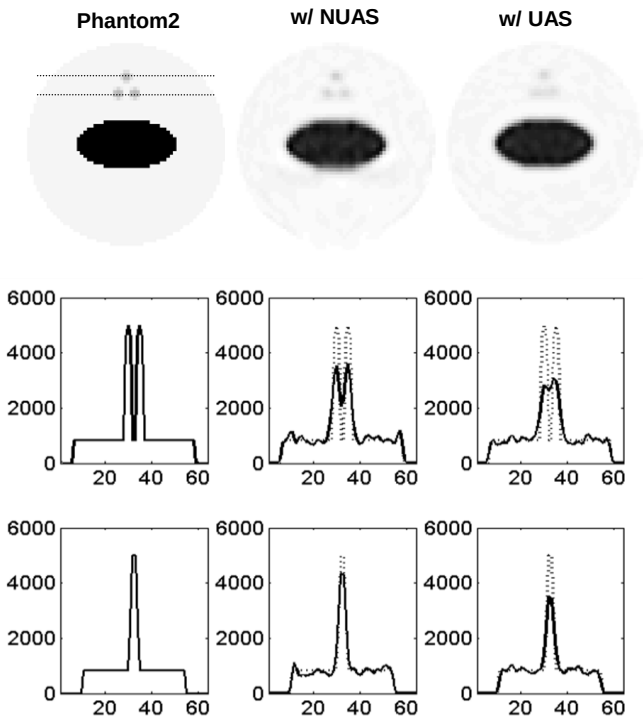

Figure 5.9 Left column: 2-D and 1-D cross sections of the simulated phantom2; Central column: cross sections of the reconstructed phantom image obtained with non-uniform angular sampling (NUAS); Right column: image obtained with uniform angular sampling (UAS). The x-axis is the pixel number across the object-space

The optimum time distribution obtained with the proposed approach is a complex function of many factors, including the location of the ROI with respect to the orbit of the camera, the relative location and strength of the background features (such as the elliptical features shown in Figure 5.7 and Figure 5.8, the attenuation of the gamma ray originated from the ROI, etc). While the effects of these factors are coupled to each other and therefore difficult to quantify, one may draw a few general guidelines from these results:

- Sampling angles corresponding to larger probabilities for detecting the gamma rays originated from the ROI are generally more important and therefore associated with longer imaging time. This applies to cases of PO1-PO6 in Figure 5.7. Similarly, sampling angles that lead to greater sensitivity to strong background features are generally less favorable. This is also evident in cases of PC1-PC6 in Figure 5.7.
- Sampling angles that lead to overlapping between the projections of the ROI and other strong background features could be more favorable angles. This is evident by examining cases of PC1-PC6 in Figure 5.7. This result could be explained by considering two factors, which are balancing each other. On the one hand, such overlapping reduces the weighted correlation between the response of the detector to a target pixel and the response to other

source pixels nearby. Since this correlation is directly related to the physical resolution of the system, the reduction in correlation could help to separate the target voxel from other voxels during reconstruction. On the other hand, the overlapping between the projections tends to reduce the amount of information carried by each detected photon, since it is associated with a larger statistical uncertainty. These two effects are counter-balancing each other, which leads to these counterintuitive results.

- In Figure 5.7, we have also observed some cases that zero imaging times interleave between positive imaging times. These time distributions were carefully verified to minimize imaging variance. In contrast, a smoother time distribution seems to be attainable with the increasing number of possible sampling angles, as shown in Figure 5.8. This indicates that the irregularity is likely to come from the discretization of sampling angles and object-space.

5.4 Summary

We have developed an analytical approach for finding the optimum system configuration that minimizes image variance. This approach has three key components: (a) a set of equations for evaluating image variance and resolution properties attainable with a given sampling strategy, (b) an iterative algorithm for searching through the parameter space and finding the optimum sampling strategy and (c) a novel computation approach based on the non-uniform object-space pixelation method for reducing the computation load [94]. This combination leads to a highly efficient computation approach that can be implemented in near real-time.

In this study, we have utilized the proposed computation approach for adaptive angular sampling in SPECT imaging. It is shown that the performance of a SPECT system can be improved with non-uniform allocations of imaging time across all possible sampling angles. We have also demonstrated, with a series of Monte Carlo studies, that the proposed computation approach is highly effective for finding the optimum sampling strategy that leads to a significantly lowered image variance.

It is worth noting that the proposed adaptive angular sampling approach lays the foundation of further optimization of SPECT systems with respect to a wide range of imaging parameters. For an adaptive imaging system that offers a finite number of possible configurations, the proposed method could help to rank the relative importance of each configuration. By identifying the optimum combination of different imaging configurations in near real-time, the proposed computation approach allows the system to offer an optimum efficiency for collecting imaging information about unknown objects. This makes the proposed computation approach well suited for future adaptive nuclear imaging applications. Moreover, this method could be equally

applicable to other imaging modalities, such positron emission tomography (PET) and X-ray computed tomography (CT).

CHAPTER 6

SPECT SYSTEM OPTIMIZATION AGAINST A DISCRETE PARAMETER SPACE

As previously discussed, for SPECT to become a practical tool for *in vivo* studies at ultra-high spatial resolutions, technical approaches to improve the information collection process are in critical need. This problem could be partially alleviated with the adaptive SPECT approach. In an adaptive SPECT system, the data acquisition hardware could be varied in real-time in response to the information being collected during an imaging study. This helps to offer an improved efficiency for collecting useful imaging information and therefore to provide an optimum performance for imaging an unknown object. In general, the adaptive imaging requires a practical adaptation rule to conduct the variation of the hardware for the best performance, which could be equivalently considered as the system optimization with respect to the related system parameters.

We have previously reported that both the vector MUCRB approach and the NUOP approach could be combined together to enable the almost real-time evaluation of system performance, quantified by tradeoffs between resolution/bias and variance. Based on them, an adaptive angular sampling approach is proposed for a SPECT system with one ultra-high resolution camera-head rotating around the object [50]. This method allows one to identify the optimum distribution of imaging time among all possible sampling angles for achieving the lowest image variance at a given resolution.

A further task is to optimize a SPECT imaging system against a wide range of system and imaging parameters, e.g., the number and the size of the pinholes in apertures, locations of apertures in the system, magnifications, imaging time distributions and so on. Although we are using SPECT instrument as a platform to demonstrate this system optimization strategy, the method developed in this work provides a general framework for system optimization across multiple imaging modalities, such as positron emission tomography (PET), X-ray computed tomography (CT).

6.1 Fundamental Challenges for Direct System Optimization

Suppose a SPECT imaging system is characterized by a total of S system parameters, $\boldsymbol{q}=[q_1, q_2, \ldots, q_S]$. These parameters could include detector and aperture positions, detector intrinsic resolution, configurations of pinholes on each aperture and many more. One could consider to

optimize the system, so that it gives the best performance for an imaging task, measured by a given performance index Ω,

$$\boldsymbol{q}_{optimum} == \underset{\boldsymbol{q}=[q_1,q_1,\cdots q_S]}{\operatorname{argmax}} [\Omega(\boldsymbol{q})]. \tag{6.1}$$

The system performance index Ω is typically a scalar or a vector function of $\boldsymbol{q}$. It is chosen to represent the effectiveness of the given imaging system to fulfill an imaging task or a class of tasks. For example, one could choose to optimize an imaging system for estimation tasks, based on performance indices such as spatial resolution [64, 69], image variance [58, 70] and their tradeoffs [71-74], and/or the accuracy for ROI quantification [60, 66]. Alternatively, one could optimize a system for detection tasks based on the receiver-operation characteristics (ROC) of computer or human observers [75-82].

Unfortunately, there are several fundamental challenges for direct system optimization procedure:

- To optimize a system, one needs to define an imaging task, along with one or more quantitative indices measuring the performance of any possible system configuration for the task. However, for typical SPECT systems, the relationships between these performance indices Ω and system parameters $\boldsymbol{q}$ are highly complex and often mathematically intractable, which may prohibit the use of (6.1) for direct system optimization.
- The system parameters to be optimized may vary continuously and thus a system optimization procedure typically requires a search through a vast (often infinite) number of possible system and sampling configurations. The amount of computation required makes it impractical to use brute-force approaches for finding the best system configurations.
- Equation (6.1) is incomplete for system optimization. During an imaging study, it could be sometimes beneficial to use more than one system configurations to obtain complementary imaging information. This information could be combined within the maximum likelihood estimation framework for a better image quality than those obtainable with any single system configuration alone. However, it would be difficult to accommodate this scenario with the optimization scheme based on (6.1). Moreover, combinations of multiple possible system configurations would expand the parameter space, which further excludes the use of brute-force approaches in system optimization.
- Adaptive imaging procedures [48] would require many cycles of system optimization to be completed in real-time. This further adds to the need for an efficient computation procedure for system optimization.

6.2 An Indirect Approach for SPECT System Optimization

To alleviate these difficulties, we have developed an analytical approach for optimizing SPECT system design or the sampling strategy for use with a variable imaging hardware. A SPECT imaging system could be described in the following way:

1) We consider a SPECT imaging system that consists of L sub-detection-systems. For example, a non-rotating SPECT system may have 16 camera-heads working simultaneously (hence L=16).
2) Parameter space is discretized. Suppose each sub-detection-system has a finite number of possible configurations: Sub-system 1 has l_1 options, Sub-system 2 has l_2 options and so on, and $\sum l_i = K$. For example, a SPECT has L camera-heads, each having a detector coupled to an aperture. In the parameter discretization, each pinhole aperture cannot be freely moved any more but only placed at one of M different locations. This gives K=L ×M. In practice, some parameters may be meaningful to be discretized. For example, it is not necessary to continuously vary pinhole size, but only need to attempt some possibilities, such as 200 μm, 250 μm and so on.
3) Each sub-detection-system with a unique configuration is defined as a virtual detector. The above-mentioned SPECT system could be considered as having K virtual detectors.
4) For an imaging study we setup an artificial scenario, in which we allow that all K virtual detectors could be used sequentially for imaging the object, and each is assigned with a finite imaging time $t^{(k)}$. The distribution of imaging times across all virtual detectors is constrained by the actual imaging protocol. For the foregoing SPECT system, the total imaging time assigned to all virtual detectors associated with any given camera-head should be a constant T.

With these notations, the problem of optimizing the corresponding system design or sampling strategy is turned into the task of finding the best time distribution among all the virtual detectors that are allowed to use during an imaging study.

$$\boldsymbol{t}_{opt} \equiv \left[t^{(1)}{}_{opt}, t^{(2)}{}_{opt}, \cdots, t^{(K)}{}_{opt}\right] = \underset{\substack{\boldsymbol{t}=\left[t^{(1)},\ t^{(2)},\cdots,\ t^{(K)}\right], \\ \textit{under certain constraints}}}{\arg\max} \left[\Omega(\boldsymbol{t})\right]. \tag{6.2}$$

This approach offers several interesting features. First, the proposed approach is capable of accommodating system optimization problems against any system parameter in a discretized space. Second, the optimum system configuration or the optimum combination of system configurations is found by maximizing Ω against imaging times $\boldsymbol{t}$, rather than directly against system parameters $\boldsymbol{q}$. Mathematically, this could greatly simplify the computation procedure for

system optimization. We could use adaptive angular sampling approach to realize the imaging time distribution optimization. Finally, with the procedure (6.2), the optimum sampling strategy is not limited to the use of a single system configuration. It could also be the one using a combination of multiple sampling configurations. In either case, the proposed procedure helps to identify the most effective sampling strategy by distributing non-zero imaging times among the most useful virtual detectors.

For a SPECT system equipped with an imaging hardware adjustable during an imaging study, this approach could identify the best combination of different system configurations to be used sequentially to obtain the optimum imaging performance. For designing a stationary SPECT system, the configuration that was assigned with the largest share of imaging time may be also identified as the optimum system configuration.

In this study, the system performance is quantified by resolution-variance tradeoffs. The imaging time distribution for the optimum balance could be derived by the gradient-based iterative search algorithm used in the adaptive angular sampling approach [50], detailed in Figure 5.2. In this case, superscript (k) is denoted as virtual detectors, instead of sampling angles. In the optimization process, the update of imaging times should satisfy the particular time constraints corresponding to the optimization task. In comparison, another two constrained minimization algorithms, *sqp* and *interior-point* algorithms, would be implemented by the routine <fmincon.m> in the matlab optimization toolbox [103]. It allows one to incorporate certain constraints in the minimization process.

6.3 Simulation Results

6.3.1 Introduction to a Simulated SPECT System and Phantom

We have conducted a series of simulation studies to evaluate the proposed system optimization approach. In this study, we simulated a non-rotating SPECT system that consists of 16 camera-heads arranged in a closely packed ring with diameter of 11.76 cm (distance between the front-surfaces of two opposite detectors). The detector used has 128×128 pixels of 0.175 mm × 0.175 mm in size, and each detector is coupled to a collimator with a single pinhole of 300 μm diameter. The simulated SPECT system has a variable aperture system – during an imaging study, each pinhole could be independently moved between the corresponding detector and the object to offer projections with different magnification ratios and angular coverage. The maximum aperture-to-center distance is 39.4 mm to ensure the entire object to be fully projected onto the detector. The minimum distance is 18.6 mm, in which case, the pinhole is pushed as close to the object as practically possible without touching the object. This offers the largest magnification for

a spherical region of 10 mm diameter, located in the center of the field-of-view (FOV). During an imaging study, the location of pinhole is discretized into eight possibilities equally spaced between these two extreme positions.

The object-space was divided into 127×127×127 voxels with 0.24 mm ×0.24 mm ×0.24 mm in size. The object has a spherical volume of 30 mm diameter with uniform tracer uptake. We have carried out several similar studies, with the target voxels located at 0 mm, 4.8 mm and 9.6 mm away from the center of the object respectively. The system and phantom are schematically represented in Figure 6.1.

To utilize the proposed system optimization approach, we assigned each detector with a pinhole at each of 8 possible positions as a virtual detector. This leads to a total of 144 (16×8) virtual detectors for use in an imaging study. Since the simulated SPECT system has 16 detectors operating at the same time, the total time spent at each detector, shared by all eight pinhole locations, is equal to a constant *T*. In this study, *T* is 480 seconds.

To further reduce the amount of computation involved, we have applied the non-uniform object-space pixelation scheme that we have previously developed in [94]. This approach allows one to represent the object-space with a dramatically reduced overall voxel count and therefore to reduce the amount of computation involved in image reconstruction and in evaluating the vector

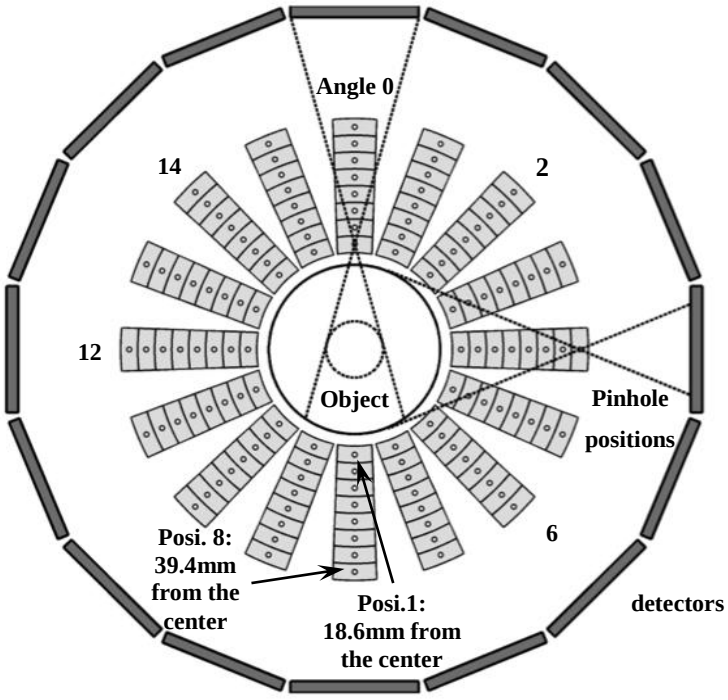

Figure 6.1 Simulated system and object geometry. Each detector is coupled to a single pinhole that can be placed at one of eight position.

modified uniform Cramer-Rao bound. In this previous work, we have demonstrated the use of the NUOP approach could preserve the image quality in regions-of-interest arbitrarily specified by the user, given that the rebinning process does not combine pixels with dramatically different values [94]. In practice, this condition could be met approximately by using the information from a short preliminary imaging study. We have also studied the impact of the non-uniform pixel rebinning on the evaluation of vector modified uniform Cramer-Rao bounds (MUCRB), and demonstrated that the rebinning process will not necessarily lead to significant errors in the system performance predicted with vector MUCRB. With this approach, the object-space were divided into voxels of different sizes, according to their distance from the center of the selected point-of-interest: a voxel size of 0.24 mm × 0.24 mm × 0.24 mm was used in a spherical region of 4 mm diameter centered at the point-of-interest, a voxel size of 0.96 mm × 0.96 mm × 0.96 mm was used in the region that is more than 2 mm but less than 5 mm away from the target-voxel. For the region that is more than 5 mm away from the target voxel, the effective voxel size was increased to 3.84 mm × 3.84 mm ×3.84 mm. With the rebinning process outline above, the total number of unknowns was shrunk to around 8000 (from about 2 million) and thus the computation load was greatly reduced. Therefore, it was possible to perform the optimization in real-time in a PC-based parallel computing environment.

6.3.2 Evaluation of the Indirect Optimization Approach

In the indirect optimization approach, the problem is simplified to minimization of performance index with respect to imaging time subject to certain time constraints. In this study, the system performance is quantified by variance of voxel centered in the object at a given resolution. In order to solve the constrained minimization problem, we have implemented three algorithms: gradient-based iterative search algorithm, detailed in Figure 5.2, and Matlab-sourced *sqp* and *interior-point* algorithms. Since these algorithms required different amount of computation per iteration, we plotted the resultant image variances against the computation time taken by a single CPU to loop through multiple iterations for fair comparison, shown in Figure 6.2. A sufficiently large number of iterations were allowed for each algorithm to ensure convergence. The figure showed that all three algorithms compared produced virtually identical final variance value once converged.

The convergence rate for the imaging time distribution was much slower than the convergence rate for the resultant image variance, shown in Figure 6.3. In fact, it took only a few iterations to achieve 90% of variance reduction (the amount of reduction is 90% of the total reduction

attainable). However, the imaging time distribution would keep changing for the next several hundreds of iterations. In fact, any time distribution results with more than 200 iterations led to virtually identical image variance and could therefore be regarded as an "optimum sampling strategy". Furthermore, with an extremely large number of iterations, both gradient-based iterative search algorithm and matlab-sourced *sqp* algorithm could provide almost the same optimum imaging time, shown in Figure 6.4.

In this simulation, all the algorithms compared had rapid initial convergence – the imaging times leading to 90% of the total possible variance reduction were found with less than 10 iterations. In reality, achieving the 90% variance reduction would probably suffice for most of system optimization task. Therefore, the proposed algorithm could offer a rapid computation approach to solve the constrained variance minimization problem and to identify the optimum system design or sampling strategy for SPECT imaging.

Given the complexity of this optimization problem, the existence of a uniquely optimum (variance-minimizing) image time distribution is not guaranteed. Even if a globally optimum imaging time distribution does exist, it is difficult to ensure that any given algorithm would not be

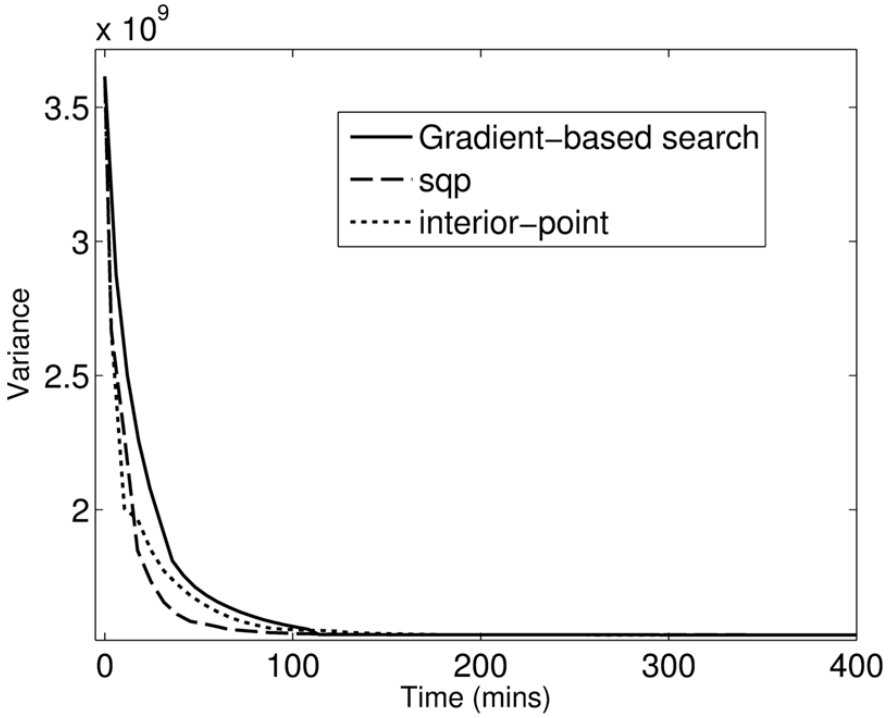

Figure 6.2 Top Upper panel: The reduction of pixel-wise variance as a function of the computation time taken by a single CPU to loop through multiple iterations with three different algorithms. The iterative approach progressively updates the imaging time distribution and leads to a monotonically reducing imaging variance values.

stuck around a local minimum and therefore misses the true variance-minimizing imaging times. However, the gradient-based iterative algorithm was demonstrated in [50], detailed in the previous chapter. In the current study, the agreement of minimized variance and optimum imaging time distribution between gradient-based iterative algorithm and matlab-sourced algorithms further validates the effectiveness of the proposed method. Therefore, even the global minimum may not be reached; at least the indirect optimization method provides an efficient way to significantly reduce the variance.

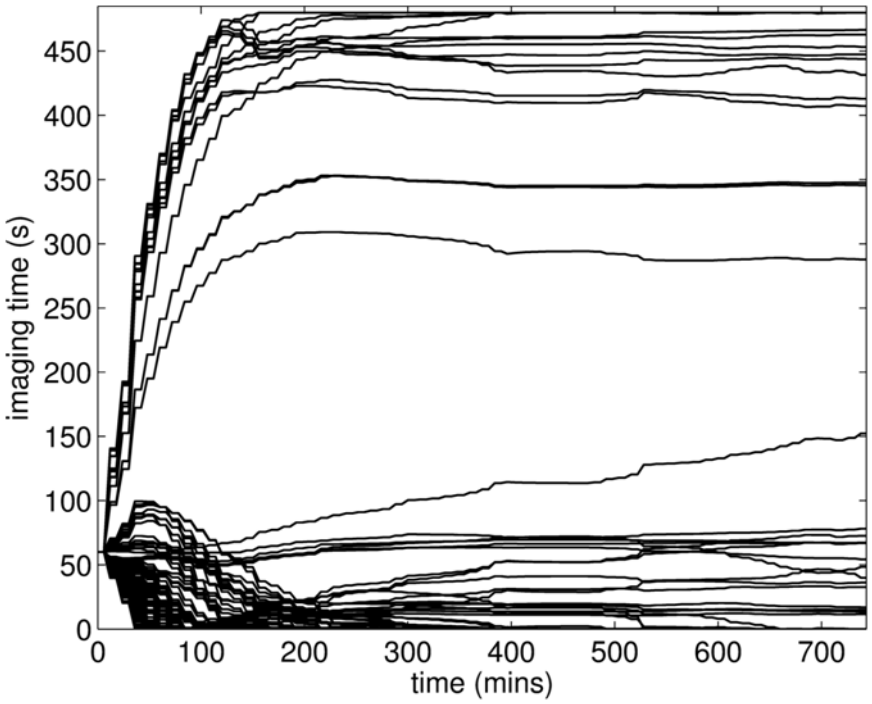

Figure 6.3 optimized imaging times for 144 virtual detectors versus calculation time curves. The convergence speed is much slower than that of variance.

6.3.3 Optimization of Sampling Strategy for Imaging Different Target Regions using Variable SPECT

In order to further demonstrate the effectiveness of the proposed system optimization approach, we have carried out a series of studies based on the simulated SPECT imaging system, seen in Figure 6.1. In these studies, we have defined three target voxels located at the center, 4.8 mm away from the center and 9.6 mm away from the center of the object respectively. For this variable-geometry SPECT system, the optimum sampling schemes that deliver the minimum

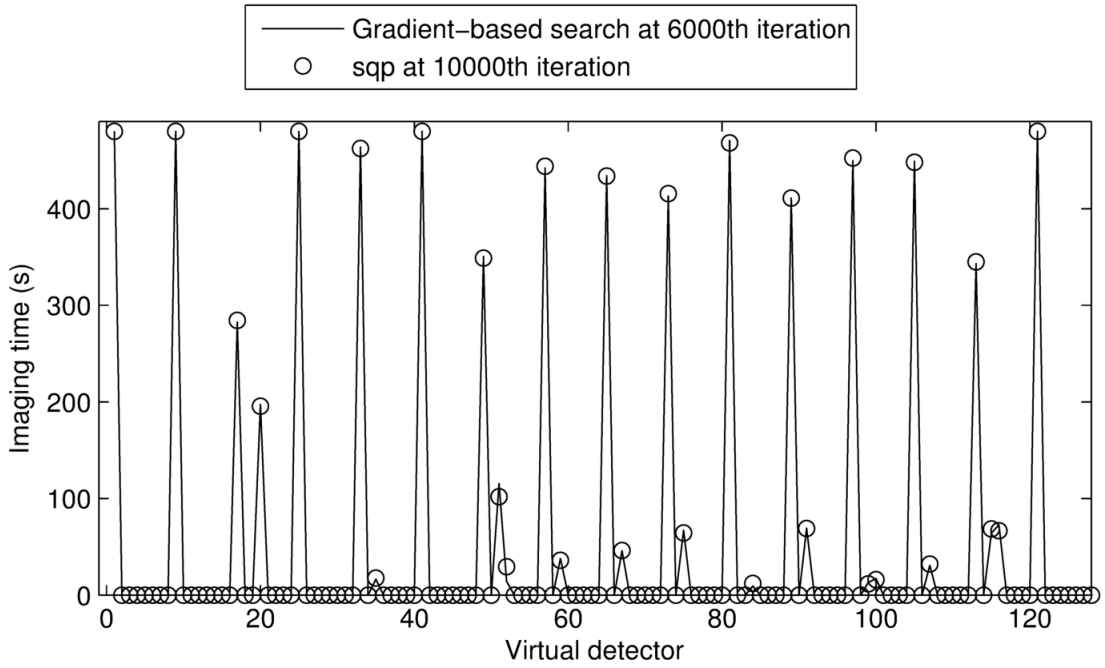

Figure 6.4 Optimum imaging times obtained by gradient-based iterative search algorithm and matlab-sourced *sqp* algorithm. Extremely large iterations are used to guarantee the time convergence. These two methods produce almost identical final results.

image variance at the target pixels were derived using the proposed indirect system optimization process.

Figure 6.5 shows the imaging time distributions as identified using the gradient-based iterative search algorithm after different number of iterations. In this case, the voxel-of-interest is at the center of the object, and the starting guess is that the total imaging time T (480 seconds) is equally distributed through all 8 possible aperture positions. As shown in this figure, the imaging time is indeed converging towards the expectation – all pinholes should be placed as close to the object as possible and remaining there throughout the entire data acquisition period. Intuitively, this geometry provides the maximum sensitivity and magnification ratio and therefore delivers the best imaging performance for the target pixel.

This study has led to a few other interesting results. First, since the point-of-interest is slightly off the center (by the same 0.5 voxel on X, Y and Z directions), the resultant imaging time distribution is not symmetric across all angles. Secondly, for the slightly off-centered point-of-interest, the optimized imaging times include a small certain fraction of imaging time for several pinholes to be placed further away from the object, as detailed in Table 6.1**Error! Reference ource not found.**. We have numerically verified that the corresponding image variance (1.548×10^9) is indeed smaller than the one obtained with the "conceptually ideal" imaging time distribution – all pinholes staying at the positions closest to the object at all time (1.584×10^9).

The use of multiple pinhole locations at certain angles could provide richer angular sampling for the target-voxel and other adjacent voxels, and therefore offer more favorable resolution-variance tradeoffs.

To further verify the proposed approach, we have moved the target voxel away from the center by 4.8 mm and 9.6 mm, and used the same optimization routine to obtain the optimum imaging time distribution. These results are shown in Figure 6.6 and Figure 6.7, and in Table 6.2 and Table 6.3. Under these situations, the optimization routine would indicate that one should move the pinholes away from the object at several angles, which in general agrees with the expectation that these specific detector/pinhole combinations should at least be able to "see" the target voxel, while being placed as close to the object as possible.

These simulation results show that, even for the object-completely-known situation, the optimum system performance has been achieved with a combination of multiple sampling geometries. We have verified these results by examining the actual variance values obtained using the combined sampling geometry and using individual fixed sampling geometries. These

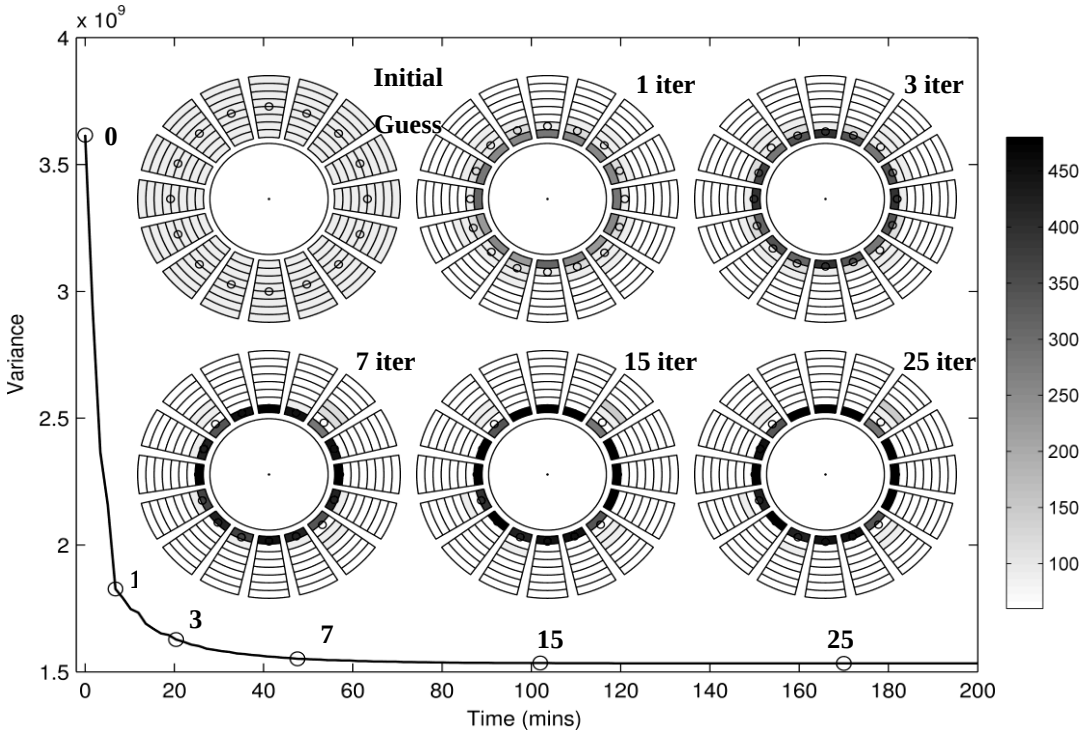

Figure 6.5 Image time distribution obtained as a function of the computation time. The target pixel is at the center of the object, as indicated with the black dot at the center. The open circles in the pie-charts represent "average" pinhole positions weighted by the imaging time distributions at individual angles. The solid circles outline the boundary of the object. The pie-charts around the objects show the image time assigned for a given pinhole location. The pie-chart at the lower right corner represents the "optimum" imaging time distribution.

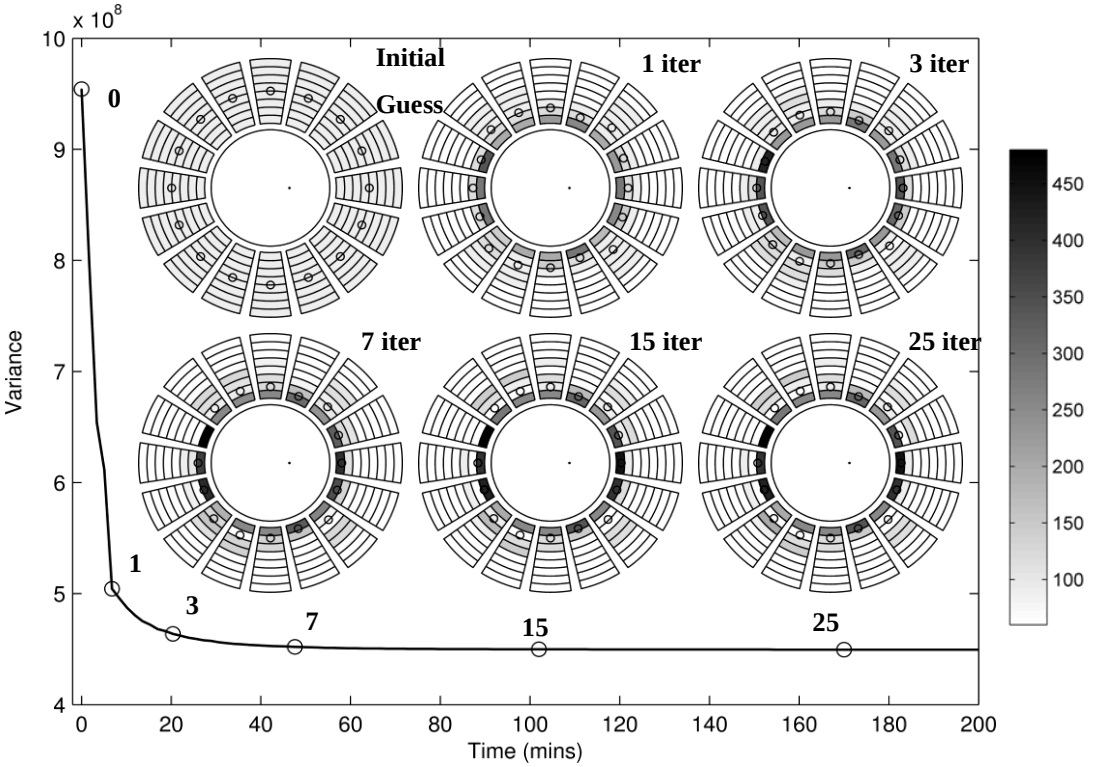

Figure 6.6 Image time distribution obtained as a function of the computation time. The target pixel is 4.8 mm off the center of the object, as indicated with the black dot.

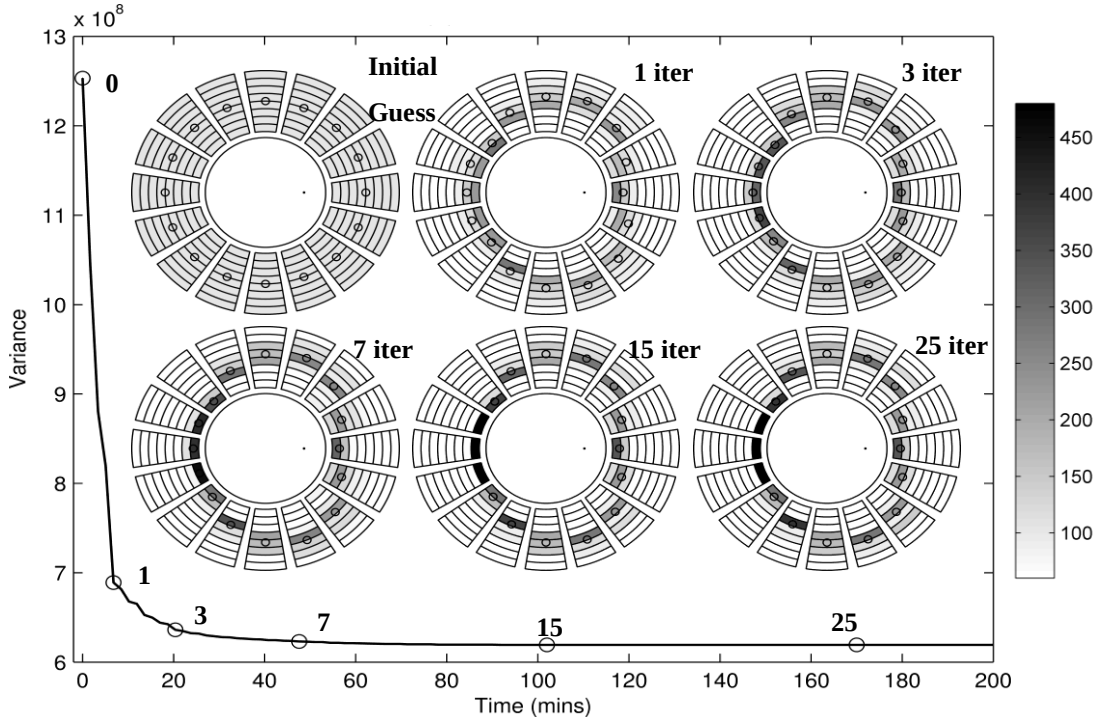

Figure 6.7 Image time distribution obtained as a function of the computation time. The target pixel is 9.6 mm off the center of the object, as indicated with the black dot.

results indicated that it is beneficial to use a synthetic sampling scheme – acquiring projection

data with multiple sampling geometries systematically chosen for a given imaging task. This allows one to acquire multiple pieces of complementary imaging information to obtain the optimum imaging performance. The analytical formula and computation approach developed in this work is naturally designed for finding the optimum synthetic sampling strategy.

Table 6.1 Optimized Imaging Time Distribution for Voxel-of-Interest at the Center of the Object

angle/position	1	2	3	4	5	6	7	8
0	480[1]	0	0	0	0	0	0	0
1	480	0	0	0	0	0	0	0
2	284.41	0	0	195.59	0	0	0	0
3	480	0	0	0	0	0	0	0
4	462.16	0	17.84	0	0	0	0	0
5	480	0	0	0	0	0	0	0
6	349.74	0	111.78	18.47	0	0	0	0
7	443.72	0	36.28	0	0	0	0	0
8	433.69	0	46.31	0	0	0	0	0
9	415.45	0	64.55	0	0	0	0	0
10	468.18	0	0	11.82	0	0	0	0
11	411.07	0	68.93	0	0	0	0	0
12	452.31	0	11.56	16.14	0	0	0	0
13	448.1	0	31.9	0	0	0	0	0
14	343.35	0	61.62	75.03	0	0	0	0
15	480	0	0	0	0	0	0	0

1. All imaging times are given in seconds.

Table 6.2 Optimized Imaging Time Distribution for Voxel-of-Interest 4.8mm off the Center of the Object

angle/position	1	2	3	4	5	6	7	8
0	294.14	79.53	71.37	0	34.95	0	0	0
1	308.03	133.47	38.49	0	0	0	0	0
2	246.39	71.18	14.93	130.14	17.35	0	0	0
3	358.6	47.96	73.44	0	0	0	0	0
4	419.58	60.42	0	0	0	0	0	0
5	372.99	94.51	12.5	0	0	0	0	0
6	262.34	37.13	0	89.54	90.99	0	0	0
7	337.12	142.88	0	0	0	0	0	0
8	269.62	83.75	98.8	0	27.83	0	0	0
9	263.24	0	65.76	151.01	0	0	0	0
10	202.35	0	277.65	0	0	0	0	0
11	410.1	0	69.9	0	0	0	0	0
12	419.22	0	60.78	0	0	0	0	0
13	480	0	0	0	0	0	0	0
14	278.53	0	90.84	110.63	0	0	0	0
15	248.05	55.43	56.26	120.26	0	0	0	0

Table 6.3 Optimized Imaging Time Distribution for Voxel-of-Interest 9.6mm off the Center of the Object

angle/position	1	2	3	4	5	6	7	8
0	0	0	0	195.52	129.88	154.6	0	0
1	0	0	0	56.45	286.39	113.75	23.4	0
2	32.75	19.58	44.19	259.45	27.94	96.09	0	0
3	90.51	236.47	72.91	63.8	16.31	0	0	0
4	337.78	129.33	12.89	0	0	0	0	0
5	88.34	237.47	88.44	37.57	28.19	0	0	0
6	30.61	0.36	56.11	226.2	44.63	122.09	0	0
7	0	0	0	23.53	296.04	100.39	60.04	0
8	0	0	0	201.81	136.63	141.56	0	0
9	0	0	384.84	54.33	40.83	0	0	0
10	295.35	141.52	43.13	0	0	0	0	0
11	480	0	0	0	0	0	0	0
12	480	0	0	0	0	0	0	0
13	480	0	0	0	0	0	0	0
14	384.9	94.45	0.65	0	0	0	0	0
15	0	0	334.63	81.13	64.24	0	0	0

6.3.4 Optimization of Sampling Strategy for Imaging Different Target Regions using Static SPECT

As most imaging systems do not allow for the configuration to be varied during the data acquisition, in this section, we have used a simple simulation study to demonstrate the potential of the proposed approach, which combines a discretized parameter space and gradient-based iterative search algorithms, for finding an optimum static SPECT system configuration that minimizes the image variance. In essence, this is a special case of the system optimization problem characterized by (6.2), but with added constraints, e.g. the imaging time could be only distributed at one of several possible configurations. Therefore, as long as one can build the constraints into the iterative search algorithm, it is possible to use a similar procedure to find the optimum static system configuration.

We have simulated a simple SPECT system as sketched in Figure 6.8. It has essentially the same geometry as shown in Figure 6.1, expect with only four detectors placed 90 degree apart. Each detector is coupled to a pinhole that can be placed at one of eight possible positions and

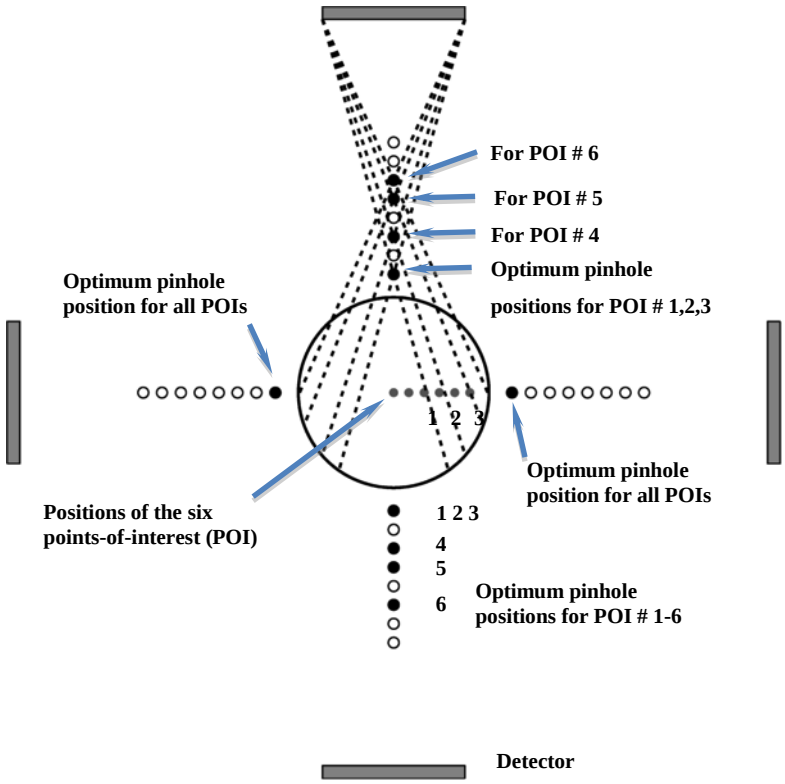

Figure 6.8 Optimum static system configurations identified for six points-of-interest, with both the gradient-based iterative search algorithm and exhaustive search. Both approaches produced exactly the same results.

remain there throughout the entire imaging study. The use of this simple geometry leads to a reduced computation load and allows us to use the brute force approach to validate the optimum system configurations identified using the method developed here.

For designing a stationary SPECT system, normally the configuration with the largest share of imaging time may be the best choice of static configuration. In practice, however, at certain angles converged imaging times may be nearly uniformly distributed across several possible pinhole positions (e.g., detector 0 and 8 in Table 6.3). In this case it is hard to determine the optimum configuration for static imaging system. To address this problem, the gradient-based iterative algorithm has been slightly modified. In the optimization process, we should update imaging time across all allowed pinhole positions at one sampling angle at one time and all other imaging times were temporarily kept as constants. After an optimum time distribution for the current sampling angle was identified, we simply chose the pinhole position with the maximum imaging time to be the "best" one and assigned all the imaging time to this position. The algorithm would then move onto the next angle, and the "best" pinhole positions at previously processed angles could be used throughout upcoming optimization process, until further updated by the angle-by-angle optimization scheme.

Figure 6.8 shows several static system configurations optimized for imaging 6 points-of-interest (POI) that are located 0 mm, 1.6 mm, 3.2 mm, 4.8 mm, 6.4 mm, and 9.6 mm from the center respectively. The changes of imaging times during optimization process at each sampling angle for observing POI #1~ #6 are shown in Figure 6.9 ~ Figure 6.14. In these figures, we used gradient-based iterative search algorithm to update imaging times at each sampling angle for 20 iterations and at the next iteration to assign all the time at the pinhole position distributed with the largest time. The imaging times at other sampling angles were fixed. Then the optimization process moved to the next sampling angles till each sampling angle had been updated three times. The figures show that in this simulation one could obtain the stable optimum time distribution for the static imaging system. In order to validate these results, we have also carried out a series exhaustive search, for each of these cases, to find the true variance-minimizing system configuration. In these six cases studied, both approaches produced exactly the same results. This confirmed that the proposed indirect optimization method has successfully identified the static system configurations that minimize the image variance.

In this example, we have used a relatively crude approach that enforces the update process to assign all the imaging time to a single pinhole position at each angle. Although it has worked well for these six cases, there could be many other ways that one can implement the same constraint with different computational procedures. The topic of building various constraints, with sufficient

numerical, computational, or theoretical efficiency depends heavily on the nature of the particular system optimization problems, which would certainly deserve further explorations.

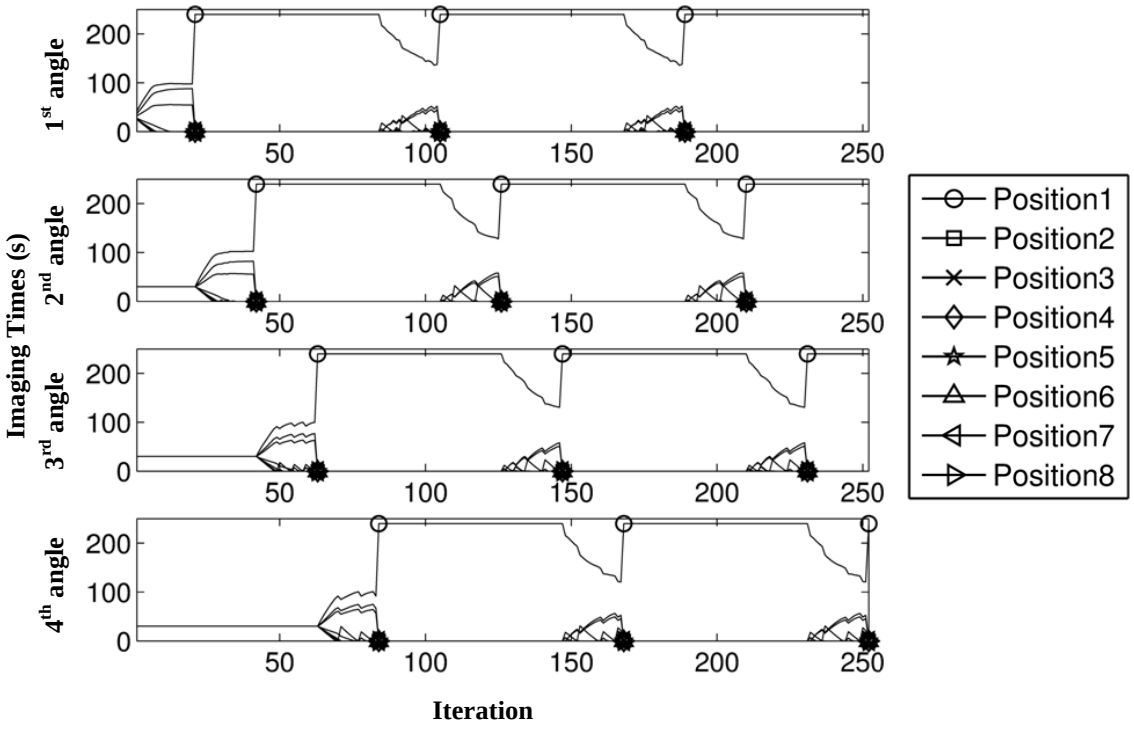

Figure 6.9 Optimized imaging time versus iteration number curves for observing POI #1 using static imaging system in Figure 6.8

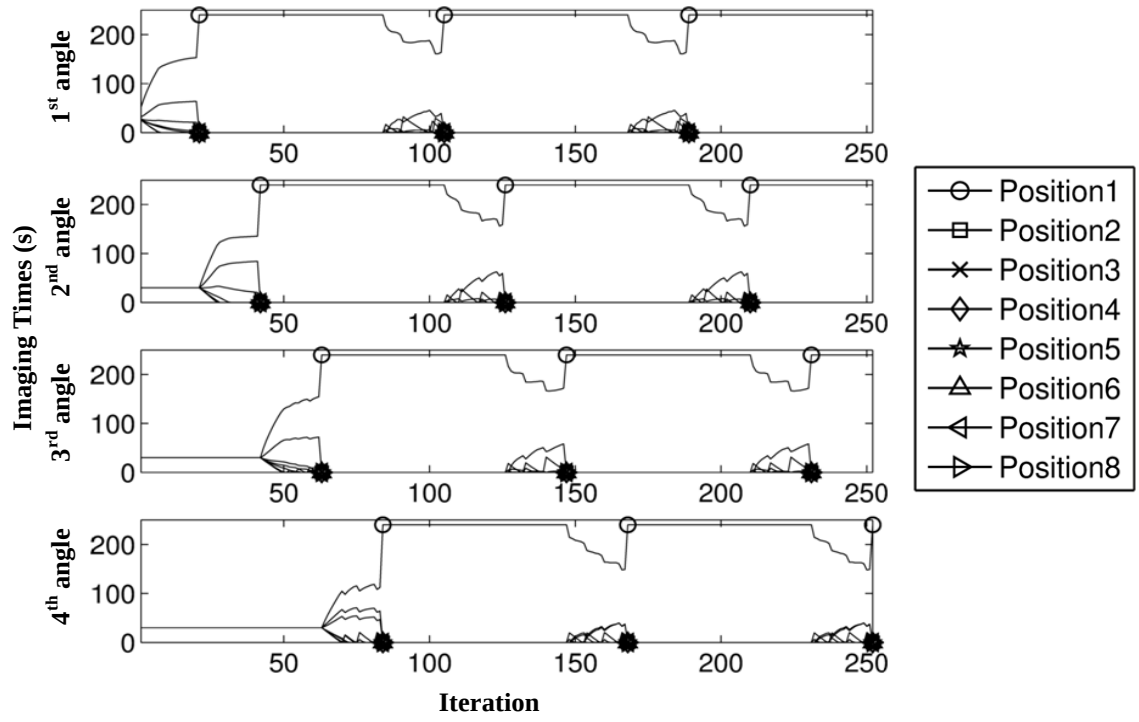

Figure 6.10 Optimized imaging time versus iteration number curves for observing POI #2

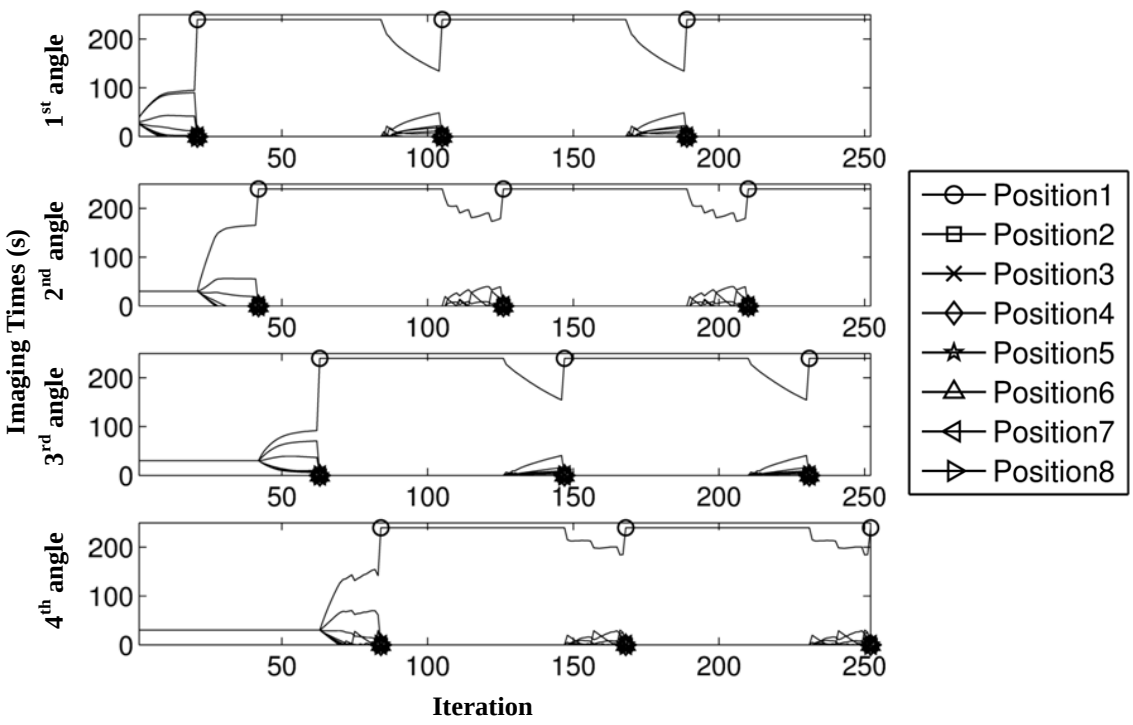

Figure 6.11 Optimized imaging time versus iteration number curves for observing POI #3

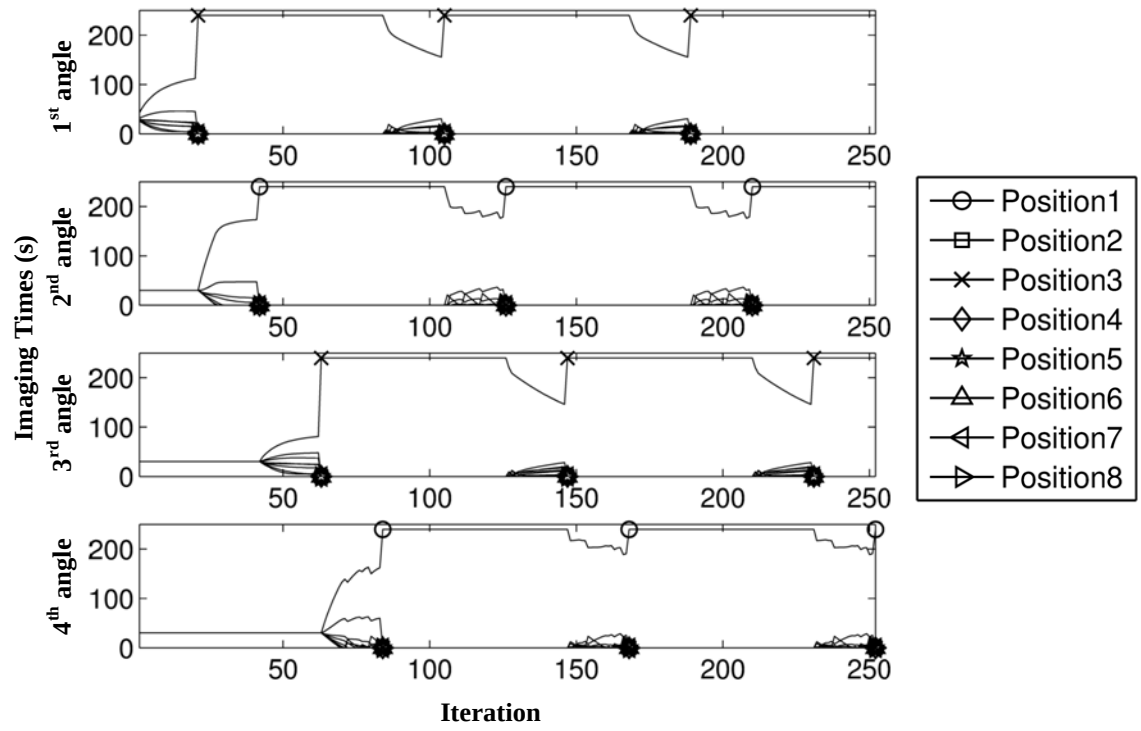

Figure 6.12 Optimized imaging time versus iteration number curves for observing POI #4

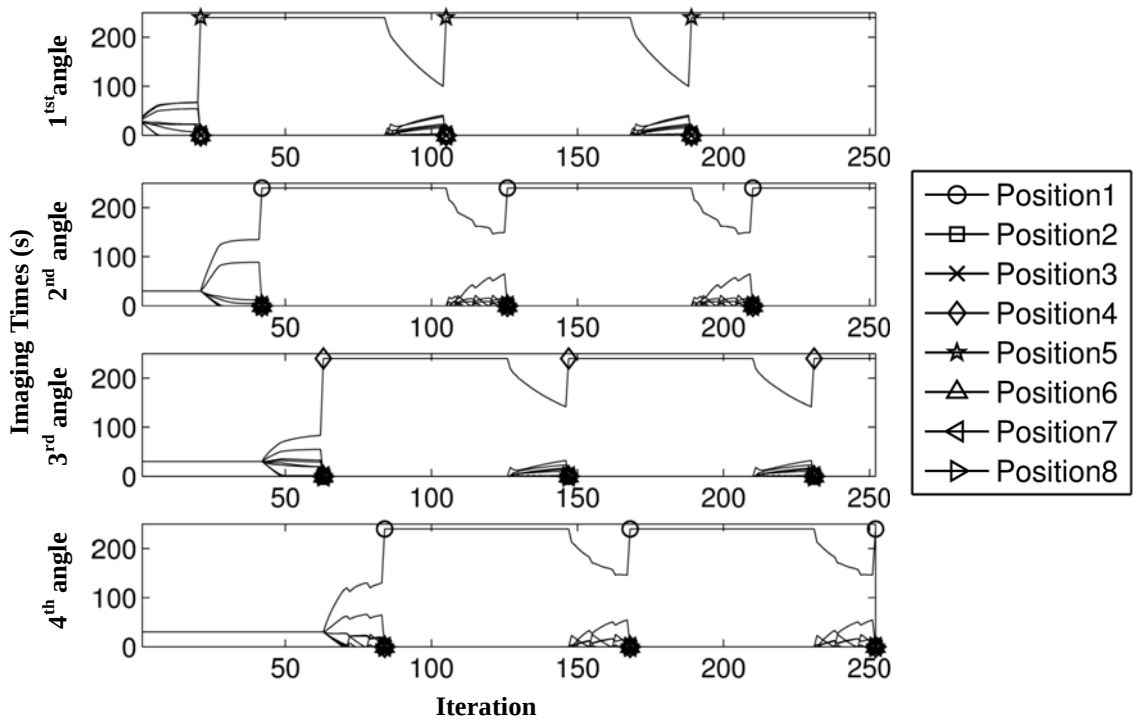

Figure 6.13 Optimized imaging time versus iteration number curves for observing POI #5

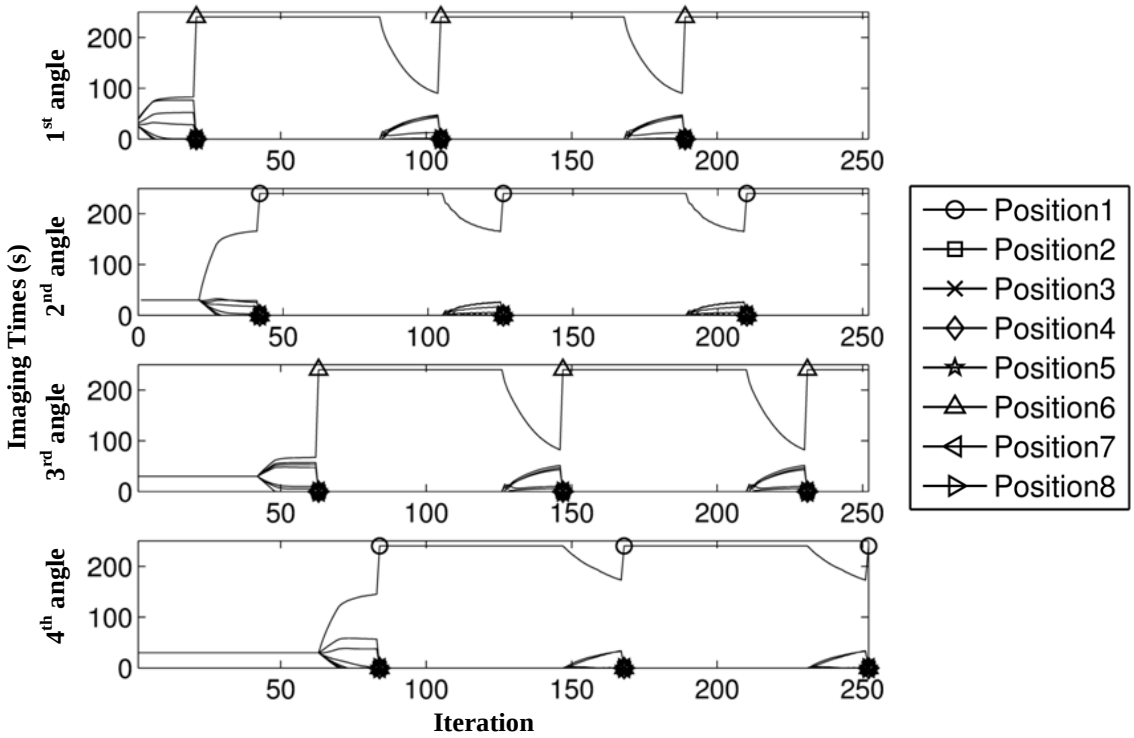

Figure 6.14 Optimized imaging time versus iteration number curves for observing POI #6

6.4 Summary

In this work, we have developed a general approach for optimizing the design of static SPECT imaging systems, or for optimizing the sampling strategy for use with a variable-geometry SPECT system. This indirect system optimization approach has several unique features. First, it offers a framework for optimizing an imaging system against almost any system parameter. Second, instead of directly optimizing performance indices against the actual parameter, this approach finds the optimum system and sampling design by (a) discretizing the parameter space and (b) optimizing the performance indices against imaging times assigned to individual possible system configurations. This approach has allowed us to derive a series of closed-form equations and an efficient computation scheme to facilitate the system optimization with a reasonable computing power.

The results presented in this work confirmed the potential benefit of using variable data acquisition hardware and a synthetic sampling strategy – multiple sampling configurations could be combined to provide an imaging performance better than those attainable with any single imaging strategy alone. The proposed approach naturally offers an interesting possibility for improving SPECT system performance for imaging unknown objects. In addition, brute-force MC simulation results also show that the optimum imaging time derived could help the design of stationary SPECT system.

Despite the reasonable success demonstrated with some simple optimization problems, we would like to highlight one of the key issues of this development. With the formulation presented, there is no proof that image variance is a concave function of the system parameters. Similarly, there is no guarantee that the proposed algorithm (or the Matlab-sourced algorithms compared) would find the global minima for image variance. One of the options for alleviating this problem is to use a multi-start approach. It repeatedly runs the search algorithm with different initial guesses systemically selected across the parameter space, and then finds the absolute minimum variance values. Further study is certainly needed to provide a better understanding of this optimization problem.

In the current study, although this method is evaluated in the simulation to optimize the locations of apertures in a non-rotating SPECT imaging system, based on the resolution-variance tradeoff criterion - an optimum system design is the one that produces the lowest possible image variance at a given spatial resolution, it is worth noting that method could be a generic optimization strategy, which is reflected in the following aspects:

- This study could be applied to optimization with respect to a wide range of system parameters as long as they could be discretized, for example, aperture configurations, detector intrinsic performance and location and so on.
- Similar as (5.9) and (5.10), the formulation and computation approaches developed here are based on a comprehensive description of the resolution and covariance properties of resultant images as functions of various system parameters. This information could be used to expand the current approach to optimize imaging systems based on other imaging tasks and performance measures, such as the bias-variance tradeoffs for region-of-interest quantification, and observer performance for detection tasks.
- This method developed could provide a general framework for system optimization across multiple imaging modalities that can be described with linear Gaussian or linear Poisson models. For example, with this approach, one could optimize the positions of a high-resolution detector in an adaptive zoom-in PET system [52] or optimize the magnification, location and shape of x-ray beam for an adaptive CT [53].

CHAPTER 7

CONCLUSIONS AND FUTURE WORK

7.1 Conclusions

In the development of SPECT system for imaging small animal *in vivo*, higher spatial resolution is continuously sought to improve the spatial detail and image quality. However, the pursuit of high resolution must result in the poor collection efficiency and limited quantitative accuracy, which ultimately confines the capacities of the imaging system. There exists a tradeoff between resolution and quantitative accuracy.

Therefore the key challenge for ultra-high resolution SPECT system design is how to optimize the system configuration, especially for observing the unknown object, with respect to a wide range of design and imaging parameters, e.g., the number and the size of the pinholes in apertures, locations of apertures in the system, magnifications, imaging time distributions and so on. Our objective is to maximize the efficiency for collecting useful imaging information regarding a given task, and therefore obtain an optimum image quality. It is worth noting that the system configuration optimization for maximizing the collection efficiency is also meaningful for other high resolution nuclear imaging modalities, such as PET or CT.

In practice, the optimization process is necessary to be realized in real-time but the computational requirements may not be compatible with complexity of acquisition parameters, even with state-of-the-art performance parallel computing platform. On the one hand, algorithms of image reconstruction and system performance evaluation are often time-consuming. On the other hand, for typical imaging systems, the relationships between figures-of-merit and system parameters are highly complex and often mathematically intractable. A generally accepted analytical approach for direct system optimization against the design parameters may not be available. Furthermore, a system optimization procedure typically requires a search through a vast (often infinite) number of possible system and sampling configurations. Thus it is almost impossible to use brute-force approaches to search through the vast parameter space.

To address these problems, we have developed four approaches. Both vector modified uniform Cramer-Rao bound and non-uniform object-space pixelation approach have been developed for rapidly evaluating system performance indices, such as resolution-variance tradeoffs and bias-variance tradeoffs. In addition, both adaptive angular sampling approach and indirect system optimization approach have been developed for efficiently optimize system

configuration and/or sampling strategy within infinite parameter space. The aim is to enable the optimization of system collection efficiency for imaging useful information about the unknown object with reasonable computation load.

The fundamental performance of an imaging system is usually quantified by the tradeoffs between spatial resolution/bias and the noise in reconstructed images. Since the commonly-used brute-force MC simulation is too time-consuming to evaluate the tradeoffs in high resolution imaging application, the vector modified uniform Cramer-Rao bound (MUCRB) approach is developed as an alternative. In this method, one could calculate the bounds on estimator variance at given resolutions, or equivalently, the optimum average resolution-variance tradeoffs over arbitrarily chosen voxels. In addition, the optimum bias-variance tradeoffs for ROI quantification could also be derived. Then it could be demonstrated that this method makes sense because the involved computation load could be greatly reduced while the optimum tradeoffs predicted by vector MUCRB could be asymptotically achieved by post-filtered penalized maximum likelihood estimators with well-defined penalty function in linear Poisson model. This method may serve as the first step towards the system optimization.

Although vector MUCRB approach could be implemented to analytically quantify the system resolution and statistical noise characteristic instead of time-consuming brute-force MC simulation, the required calculation load in high resolution imaging application is still very huge. So the NUOP approach has been developed in order to further reduce the computation effort. The basic idea behind the method is that since the major imaging task is usually to reveal the microscopic structure inside a target-region, one may tolerate lower spatial resolutions in areas outside the target-region. Therefore, the object-space could be divided into smaller voxels for target-regions, and into larger voxels in areas that are relatively smooth and/or less important to the reconstruction of the target-regions. As the key to the non-uniform object-space pixelation approach, the FIM-based and distance-based criteria have been used in determining systematic rebinning strategies. As a result, the NUOP approach could reduce the calculation load by 1-2 orders of magnitude without scarifying image quality inside the ROI. Furthermore, this method could be adapted to the vector MUCRB for rapid evaluation of the resolution-variance and bias-variance tradeoffs, which could be considered as an important step towards a feasible system performance optimization.

Because of the development of the vector MUCRB approach and the NUOP approach, it is feasible to evaluate the system performance in almost real time, which lays the foundation for practical system performance optimization. At first the optimization aspect is limited to imaging time distribution for a rotating SPECT system. In the cases that the ROI is superimposed on or

located near other strong background features, the conventional uniform angular sampling approach may not provide the most efficient way for gathering imaging information related to the ROI. So at first we have developed an adaptive angular sampling approach to optimize imaging time distribution for a rotating SPECT system to improve collection efficiency. This approach allows the camera to spend larger fractions of imaging time at angles those are relatively more important for acquiring useful imaging information regarding a given imaging task. Even variance could be mathematically expressed as a function of imaging time distribution in close-form, the computation load for searching through the entire continuous imaging time domain could be too heavy for regular PC-based computation environment. Therefore, we have developed a gradient-based search algorithm to iteratively refine the direction of the change of the imaging time for reducing images variance. Together with vector MUCRB approach and the NUOP approach, we could realize the nearly real-time optimization of imaging time distribution that delivers the minimum variance at a given resolution.

Till now we could use adaptive angular sampling approach to optimize the imaging time distribution for a rotating SPECT system. We would like to further develop an optimization approach for an arbitrarily given set of system parameters or the sampling strategies with a variable/adaptive SPECT imaging hardware. Instead of directly optimizing performance indices against the actual parameter, this approach finds the optimum system and sampling design by discretizing the parameter space and then optimizing the performance indices against imaging times assigned to individual possible system configurations. All the approaches developed could be used together to provide an efficient computation scheme to facilitate the optimization with a reasonable computing power. Moreover, it has been demonstrated that the indirect optimization method could also help the design of stationary SPECT system.

Given the complexity of this optimization problem, the existence of a uniquely optimum image time distribution is not guaranteed. Even if a globally optimum imaging time distribution does exist, it is difficult to ensure that the given algorithm would not be stuck around a local minimum and therefore misses the true variance-minimizing imaging times. However, the effectiveness of adaptive angular sampling approach is previously demonstrated in brute-force MC simulation studies. The agreement of minimized variance and optimum imaging time distribution between gradient-based iterative algorithm and Matlab-sourced algorithms further validates the effectiveness of the proposed method. Therefore, even the global minimum may not be reached; at least the indirect optimization method provides an efficient way to significantly reduce the variance.

In the current study, although we are using SPECT imaging as a platform to demonstrate this system optimization strategy relied upon the resolution-variance tradeoff criterion - an optimum system design is the one that produces the lowest possible image variance at a given spatial resolution, the general framework developed in this effort could be used to optimize the instrumentation design in other modalities, such as positron emission tomography (PET), X-ray computed tomography (CT), based on other imaging tasks and performance measures, such as the bias-variance tradeoffs for region-of-interest quantification, and observer performance for detection tasks.

In order to implement the optimization task, there is a considerable interest to utilize the adaptive imaging concept. In the adaptive data acquisition, the system hardware or acquisition protocol could be autonomously varied in real-time in response to the information being collected during an imaging study. The objective is to maximize collection efficiency for imaging an unknown object and provide the best image quality.

The development of these four approaches gives a feasible adaptation rule to optimize the system configuration in almost real-time. As the first step, full system response functions could be pre-calculated in the object-space modeled with sufficiently small and uniformly-sized voxels for all the system parameters to be optimized. This step is performed only once. And then the object could be imaged in a short time and rapidly reconstructed in an object-space with coarse and uniform pixelation. The heavily smoothed reconstruction could be regarded as a good approximation to unknown source. Based on it, we could identify the ROI and finish all the rebinning operations. The required computation load at this step could be ignored compared with that in optimization process. This preliminary information may also come from other imaging modalities, such as PET or CT. At last, the indirect system optimization approach should be realized in non-uniformly pixelated object-space to obtain the optimum configurations and corresponding imaging time distribution in nearly real-time, which will be used in the following imaging procedure to acquire experimental data from the object for the optimum image quality.

In the optimization process, it is worth noting that the required mean projection data used in Fisher information matrix, seen in (5.6) and (5.7), is usually unavailable in practice. But one can approximate it using the preliminary image by just forward projecting it into detector space. This approximation is reasonable since forward projection can be considered as a smooth operator. Furthermore, in the Fisher information matrix the diagonal matrix depends on mean projection is sandwiched between $\boldsymbol{A}^{\mathrm{T}}$ and $\boldsymbol{A}$, which induces the further smooth [52, 58, 64].

7.2 *Future Work*

In the previous theoretical study we have developed a series of approaches for practical performance evaluation and optimization for the high resolution imaging system in almost real-time. The next phase will be to build an adaptive imaging system to further examine and validate these theoretical predictions in a realistic situation.

In the development of system hardware, we may use the latest Gen-IV energy-resolved photon-counting (ERPC) CdTe detectors, seen in Figure 7.1 [104]. The dimension of the detector is 11 mm × 22 mm. The ERPC ASIC has a total of 2048 readout channels arranged in an array of 32 ×64 pixels. The pixel pitch size is 0.35 mm × 0.35 mm and the intrinsic energy resolution of the ERPC detector is around 3 keV at 140 keV events. The excellent energy resolution and high spatial resolution ensure the ERPC detector to provide promising imaging performance in adaptive SPECT system, while the compact size brings a great convenience in adaptive system design and sufficient accuracy in geometry alteration.

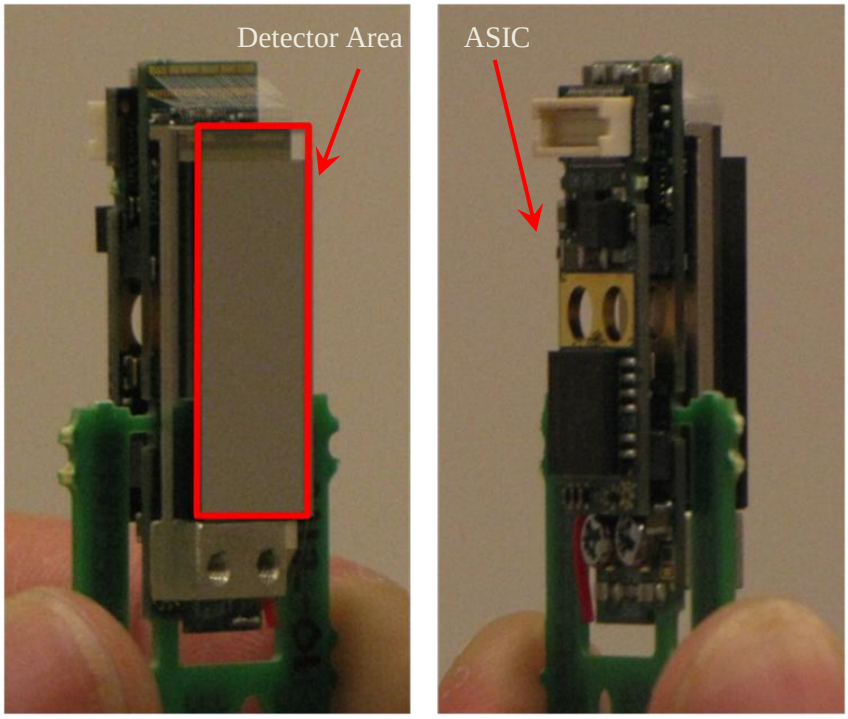

Figure 7.1 Gen-IV energy-resolved photon-counting (ERPC) CdTe detectors

In the development of adaptive imaging system, it is also desired to build a complete data acquisition and processing software system in a parallel computing environment which consists of modules of system calibration, system response functions generation, non-uniform object-space pixelation, experimental data acquisition, image reconstruction and optimization of system parameters using the indirect optimization method. As a result, we could reduce the waste of time

during data processing and transmission as much as possible and help to realize the optimization in real-time.

With the adaptive imaging system, at first we could experimentally demonstrate the effectiveness of angular sampling approach for optimizing imaging time distribution in observing the ROI, which is the basis in the indirect optimization method. And then we could explore the synthetic sampling scheme in practice – acquiring projection data with multiple sampling geometries systematically chosen for a given imaging task, which may be a future direction for SPECT imaging. Furthermore, we could realize the adaptive sampling acquisition in phantom and mouse study to verify the improvement of image quality. This could help us to better investigate the properties of the body's physiology,

Finally, current computations are implemented in a PC-based parallel computing environment with 32 CPUs built in a cluster of four workstations. Data communication between different workstations relies on the 1 Gigabit Ethernet control network. In the simulations introduced before, we could find the optimum imaging time distribution within several minutes, which results in the converged minimum variance. In the future, we could possibly transplant the computation codes in a much better computation environment, the university of Illinois shared computing cluster. This cluster is designed to support up to 512 nodes for computing applications, each configured with (2) Intel HP X5650 2.66Ghz 6C Processors. The nodes are interconnected with a 10 Gigabit Ethernet control network. The disk system was selected to support expandability and the GPFS file system [105]. Built in this outstanding parallel computing cluster, the computation involved in data transmission could be ignored and thus the computation speed could grow as a linear function of the computation nodes used. Therefore it is practical to further improve the computation speed and realize adaptive system optimization in real-time definitely.

REFERENCE

[1] M. N. Wernick, and J. N. Aarsvold, *Emission Tomography: The Fundamentals of PET and SPECT*: Elsevier Academic Press, 2004.

[2] G. F. Knoll, *Radiation Detection and Measuremen*t, John Wiley, 2000.

[3] J. Y. Qi, and R. M. Leahy, "Iterative reconstruction techniques in emission computed tomography," *Physics in Medicine and Biology,* vol. 51, no. 15, pp. R541-R578, Aug 7, 2006.

[4] J. L. Prince, and J. Links, *Medical Imaging Signals and Systems,* 1 ed.: Prentice Hall, 2005.

[5] "PET scan of the human brain," http://en.wikipedia.org/wiki/PET_scan, 2010.

[6] F. Rosch, and G. J. Beyer, "Quantification of uptake kinetics and radiation doses of radiotherapeutics labelled with rare earth radioisotopes using positron-emission-tomography," *Journal of Inorganic Biochemistry,* vol. 86, no. 1, pp. 91-91, Aug, 2001.

[7] A. Rahmim, and H. Zaidi, "PET versus SPECT: strengths, limitations and challenges," *Nuclear Medicine Communications,* vol. 29, no. 3, pp. 193-207, Mar, 2008.

[8] B. E. Hammer, N. L. Christensen, and B. G. Heil, "Use of a Magnetic-Field to Increase the Spatial-Resolution Positron Emission Tomography," *Medical Physics,* vol. 21, no. 12, pp. 1917-1920, Dec, 1994.

[9] R. R. Raylman, B. E. Hammer, and N. L. Christensen, "Combined MRI-PET scanner: A Monte Carlo evaluation of the improvements in PET resolution due to the effects of a static homogeneous magnetic field," *IEEE Transactions on Nuclear Science,* vol. 43, no. 4, pp. 2406-2412, Aug, 1996.

[10] N. L. Christensen, B. E. Hammer, B. G. Heil *et al.*, "Positron Emission Tomography within a Magnetic-Field Using Photomultiplier Tubes and Lightguides," *Physics in Medicine and Biology,* vol. 40, no. 4, pp. 691-697, Apr, 1995.

[11] A. J. Reader, S. Ally, F. Bakatselos *et al.*, "One-pass list-mode EM algorithm for high-resolution 3-D PET image reconstruction into large arrays," *IEEE Transactions on Nuclear Science,* vol. 49, no. 3, pp. 693-699, Jun, 2002.

[12] M. Rafecas, G. Boning, B. J. Pichler *et al.*, "Effect of noise in the probability matrix used for statistical reconstruction of PET data," *IEEE Transactions on Nuclear Science,* vol. 51, no. 1, pp. 149-156, Feb, 2004.

[13] S. DeBenedetti, C. E. Cowan, W. R. Konneker *et al.*, "On the Angular Distribution of Two-Photon Annihilation Radiation," *Physical Review,* vol. 77, no. 2, 1950.

[14] A. Rahmim, and H. Zaidi, "PET versus SPECT: strengths, limitations and challenges," *Nuclear Medicine Communications.* , 2008.

[15] E. J. Hoffman, S. C. Huang, D. Plummer *et al.*, "Quantitation in Positron Emission Computed-Tomography .6. Effect of Nonuniform Resolution," *Journal of Computer Assisted Tomography,* vol. 6, no. 5, pp. 987-999, 1982.

[16] E. Yoshida, Y. Kimura, K. Kitamura *et al.*, "Calibration procedure for a DOI detector of high resolution PET through a Gaussian mixture model," *IEEE Transactions on Nuclear Science,* vol. 51, no. 5, pp. 2543-2549, Oct, 2004.

[17] C. M. Kao, X. C. Pan, and C. T. Chen, "Accurate image reconstruction using DOI information and its implications for the development of compact PET systems," *IEEE Transactions on Nuclear Science,* vol. 47, no. 4, pp. 1551-1560, Aug, 2000.

[18] R. S. Miyaoka, T. K. Lewellen, J. H. Yu *et al.*, "Design of a depth of interaction (DOI) PET detector module," *IEEE Transactions on Nuclear Science,* vol. 45, no. 3, pp. 1069-1073, Jun, 1998.

[19] T. K. Lewellen, "Recent developments in PET detector technology," *Physics in Medicine and Biology,* vol. 53, no. 17, pp. R287-R317, Sep 7, 2008.

[20] H. Peng, and C. S. Levin, "Recent Developments in PET Instrumentation," *Current Pharmaceutical Biotechnology,* vol. 11, no. 6, pp. 555-571, Sep, 2010.
[21] A. Vandenbroucke, A. M. K. Foudray, P. D. Olcott *et al.*, "Performance characterization of a new high resolution PET scintillation detector," *Physics in Medicine and Biology,* vol. 55, no. 19, pp. 5895-5911, Oct 7, 2010.
[22] W. W. Moses, "Time of flight in PET revisited," *IEEE Transactions on Nuclear Science,* vol. 50, no. 5, pp. 1325-1330, Oct, 2003.
[23] Z. He, "Review of the Shockley-Ramo theorem and its application in semiconductor gamma-ray detectors," *Nuclear Instruments and Methods in Physics Research Section A,* vol. 463, no. 1-2, 2000.
[24] L. J. Meng, and Z. He, "Estimate interaction timing in a large volume HgI2 detector using cathode pulse waveforms," *Nuclear Instruments & Methods in Physics Research Section a-Accelerators Spectrometers Detectors and Associated Equipment,* vol. 545, no. 1-2, pp. 234-251, Jun 11, 2005.
[25] L. J. Meng, and Z. He, "Exploring the limiting timing resolution for large volume CZT detectors with waveform analysis," *Nuclear Instruments & Methods in Physics Research Section a-Accelerators Spectrometers Detectors and Associated Equipment,* vol. 550, no. 1-2, pp. 435-445, Sep 11, 2005.
[26] P. M. Bloomfield, S. Rajeswaran, T. J. Spinks *et al.*, "The Design and Physical Characteristics of a Small Animal Positron Emission Tomograph," *Physics in Medicine and Biology,* vol. 40, no. 6, pp. 1105-1126, Jun, 1995.
[27] J. S. Huber, and W. W. Moses, "Conceptual design of a high-sensitivity small animal PET camera with 4 pi coverage," *IEEE Transactions on Nuclear Science,* vol. 46, no. 3, pp. 498-502, Jun, 1999.
[28] T. K. Lewellen, "The Challenge of Detector Designs for PET," *American Journal of Roentgenology,* vol. 195, no. 2, pp. 301-309, Aug, 2010.
[29] S. Adak, R. Bhalla, K. K. V. Raj *et al.*, "Radiotracers for SPECT imaging: current scenario and future prospects," *Radiochimica Acta,* vol. 100, no. 2, pp. 95-107, 2012.
[30] F. J. Beekman, and B. Vastenhouw, "Design and simulation of a high-resolution stationary SPECT system for small animals," *Physics in Medicine and Biology,* vol. 49, no. 19, pp. 4579-4592, Oct 7, 2004.
[31] L. J. Meng, N. H. Clinthorne, S. Skinner *et al.*, "Design and feasibility study of a single photon emission microscope system for small animal I-125 imaging," *IEEE Transactions on Nuclear Science,* vol. 53, no. 3, pp. 1168-1178, Jun, 2006.
[32] Y. Shao, R. Yao, T. Ma *et al.*, "Initial studies of PET-SPECT dual-tracer imaging," in IEEE NSS/MIC, HAWAII,US, 2007.
[33] G. S. P. Mok, Y. C. Wang, and B. M. W. Tsui, "Quantification of the Multiplexing Effects in Multi-Pinhole Small Animal SPECT: A Simulation Study," *IEEE Transactions on Nuclear Science,* vol. 56, no. 5, pp. 2636-2643, Oct, 2009.
[34] F. van der Have, B. Vastenhouw, R. M. Ramakers *et al.*, "U-SPECT-II: An Ultra-High-Resolution Device for Molecular Small-Animal Imaging," *Journal of Nuclear Medicine,* vol. 50, no. 4, pp. 599-605, Apr 1, 2009.
[35] L. R. Furenlid, D. W. Wilson, Y. C. Chen *et al.*, "FastSPECT II: A second-generation high-resolution dynamic SPECT imager," *IEEE Transactions on Nuclear Science,* vol. 51, no. 3, pp. 631-635, Jun, 2004.
[36] L. J. Meng, J. W. Tan, K. Spartiotis *et al.*, "Preliminary evaluation of a novel energy-resolved photon-counting gamma ray detector," *Nuclear Instruments & Methods in Physics Research Section a-Accelerators Spectrometers Detectors and Associated Equipment,* vol. 604, no. 3, pp. 548-554, Jun 11, 2009.

[37] B. W. Miller, L. R. Furenlid, S. K. Moore *et al.*, "System integration of FastSPECT III, a dedicated SPECT rodent-brain imager based on BazookaSPECT detector technology," in NSS/MIC, IEEE, Orlando, FL, 2009.

[38] H. Kim, L. R. Furenlid, M. J. Crawford *et al.*, "SemiSPECT: A small-animal single-photon emission computed tomography (SPECT) imager based on eight cadmium zinc telluride (CZT) detector arrays," *Medical Physics,* vol. 33, no. 2, pp. 465-474, Feb, 2006.

[39] R. Accorsi, M. Autiero, L. Celentano *et al.*, "MediSPECT: Single photon emission computed tomography system for small field of view small animal imaging based on a CdTe hybrid pixel detector," *Nuclear Instruments & Methods in Physics Research Section a-Accelerators Spectrometers Detectors and Associated Equipment,* vol. 571, no. 1-2, pp. 44-47, Feb 1, 2007.

[40] L. J. Meng, "An intensified EMCCD camera for low energy gamma ray imaging applications," *IEEE Transactions on Nuclear Science,* vol. 53, no. 4, pp. 2376-2384, Aug, 2006.

[41] M. A. d. Reis, J. Mejia, I. R. Batista *et al.*, "Single Photon Emission Microscope (SPEM): A State-of-the-Art Instrument for High Resolution Molecular Imaging of Small Animal Organs " *Alasbimn Journal* 2011.

[42] L. J. Meng, G. Fu, E. J. Roy *et al.*, "An ultra-high resolution SPECT system for I-125 mouse brain imaging studies," *Nuclear Instruments & Methods in Physics Research Section a-Accelerators Spectrometers Detectors and Associated Equipment,* vol. 600, no. 2, pp. 498-505, Mar 1, 2009.

[43] G. A. Kastis, L. R. Furenlid, D. W. Wilson *et al.*, "Compact CT/SPECT small-animal imaging system," *IEEE Transactions on Nuclear Science,* vol. 51, no. 1, pp. 63-67, Feb, 2004.

[44] T. E. Peterson, D. W. Wilson, and H. H. Barrett, "Application of silicon strip detectors to small-animal imaging," *Nuclear Instruments & Methods in Physics Research Section a-Accelerators Spectrometers Detectors and Associated Equipment,* vol. 505, no. 1-2, pp. 608-611, Jun 1, 2003.

[45] L. A. Shepp, and Y. Vardi, "Maximum Likelihood Reconstruction for Emission Tomography," *IEEE Transactions on Medical Imaging,* vol. 1, no. 2, pp. 113-122, 1982.

[46] H. M. Hudson, and R. S. Larkin, "Accelerated Image-Reconstruction Using Ordered Subsets of Projection Data," *IEEE Transactions on Medical Imaging,* vol. 13, no. 4, pp. 601-609, Dec, 1994.

[47] H. H. Barrett, L. R. Furenlid, M. Freed *et al.*, "Adaptive SPECT," *IEEE Transactions on Medical Imaging,* vol. 27, no. 6, pp. 775-788, Jun, 2008.

[48] E. Clarkson, M. A. Kupinski, H. H. Barrett *et al.*, "A task-based approach to adaptive and multimodality imaging," *Proceedings of the IEEE,* vol. 96, no. 3, pp. 500-511, Mar, 2008.

[49] M. Freed, M. A. Kupinski, L. R. Furenlid *et al.*, "A prototype instrument for single pinhole small animal adaptive SPECT imaging," *Medical Physics,* vol. 35, no. 5, pp. 1912-1925, May, 2008.

[50] N. Li, and L. J. Meng, "Adaptive Angular Sampling for SPECT Imaging," *IEEE Transactions on Nuclear Science,* vol. 58, no. 5, pp. 2205-2218, Oct, 2011.

[51] L. Cai, N. Li, and L. J. Meng, "A Prototype Adaptive SPECT System with Self-Optimized Angular Sampling," in IEEE NSS/MIC, Valencia, Spain, 2011.

[52] J. A. Zhou, and J. Y. Qi, "Adaptive Imaging for Lesion Detection Using a Zoom-in PET System," *IEEE Transactions on Medical Imaging,* vol. 30, no. 1, pp. 119-130, Jan, 2011.

[53] J. W. Moore, H. H. Barrett, and L. R. Furenlid, "Adaptive CT for High-Resolution, Controlled-Dose,Region-of-Interest Imaging," in IEEE NSS/MIC, Orlando, FL, 2009.

[54] G. F. Knoll, *Radiation Detection and Measurement*, 2000.

[55] Harry L. Van Trees, "Detection, Estimation and Modulation Theory, Part I," John Wiley, 1968.

[56] H. H. Barrett, D. W. Wilson, and B. M. W. Tsui, "Noise Properties of the Em Algorithm .1. Theory," *Physics in Medicine and Biology*, vol. 39, no. 5, pp. 833-846, May, 1994.

[57] J. Nuyts, and J. A. Fessler, "A penalized-likelihood image reconstruction method for emission tomography, compared to postsmoothed maximum-likelihood with matched spatial resolution," *IEEE Transactions on Medical Imaging*, vol. 22, no. 9, pp. 1042-1052, Sep, 2003.

[58] J. A. Fessler, "Mean and variance of implicitly defined biased estimators (such as penalized maximum likelihood): Applications to tomography," *IEEE Transactions on Image Processing*, vol. 5, no. 3, pp. 493-506, Mar, 1996.

[59] M. A. Kupinski, and H. H. Barrett, *Small-Animal SPECT Imaging*: Springer, 2005.

[60] J. Y. Qi, and R. H. Huesman, "Theoretical study of penalized-likelihood image reconstruction for region of interest quantification," *IEEE Transactions on Medical Imaging*, vol. 25, no. 5, pp. 640-648, May, 2006.

[61] A. B. Wolbarst, "Foundations of image science," *Health Physics*, vol. 87, no. 1, pp. 93-93, Jul, 2004.

[62] D. W. Wilson, B. M. W. Tsui, and H. H. Barrett, "Noise Properties of the Em Algorithm .2. Monte-Carlo Simulations," *Physics in Medicine and Biology*, vol. 39, no. 5, pp. 847-871, May, 1994.

[63] Z. X. Cao, G. Bal, R. Accorsi *et al.*, "Optimal number of pinholes in multi-pinhole SPECT for mouse brain imaging - a simulation study," *Physics in Medicine and Biology*, vol. 50, no. 19, pp. 4609-4624, Oct 7, 2005.

[64] J. A. Fessler, and W. L. Rogers, "Spatial resolution properties of penalized-likelihood image reconstruction: Space-invariant tomographs," *IEEE Transactions on Image Processing*, vol. 5, no. 9, pp. 1346-1358, Sep, 1996.

[65] R. E. Carson, Y. C. Yan, B. Chodkowski *et al.*, "Precision and Accuracy of Regional Radioactivity Quantitation Using the Maximum-Likelihood Em Reconstruction Algorithm," *IEEE Transactions on Medical Imaging*, vol. 13, no. 3, pp. 526-537, Sep, 1994.

[66] L. Fu, J. R. Stickel, R. D. Badawi *et al.*, "Quantitative Accuracy of Penalized-Likelihood Reconstruction for ROI Activity Estimation," *IEEE Transactions on Nuclear Science*, vol. 56, no. 1, pp. 167-172, Feb, 2009.

[67] A. O. Hero, J. A. Fessler, and M. Usman, "Exploring estimator bias-variance tradeoffs using the uniform CR bound," *IEEE Transactions on Signal Processing*, vol. 44, no. 8, pp. 2026-2041, Aug, 1996.

[68] T. J. Kragh, and I. Alfred O. Hero, "Image Resolution-Variance Tradeoffs Using the Uniform Cramer-Rao Bound," http://web.eecs.umich.edu/~hero/Preprints/eusipcokragh.pdf.

[69] M. C. M. Rentmeester, F. van der Have, and F. J. Beekman, "Optimizing multi-pinhole SPECT geometries using an analytical model," *Physics in Medicine and Biology*, vol. 52, no. 9, pp. 2567-2581, May 7, 2007.

[70] J. Y. Qi, and R. M. Leahy, "Resolution and noise properties of MAP reconstruction for fully 3-D PET," *IEEE Transactions on Medical Imaging*, vol. 19, no. 5, pp. 493-506, May, 2000.

[71] L. J. Meng, and D. K. Wehe, "Feasibility study of using hybrid collimation for nuclear environmental imaging," *IEEE Transactions on Nuclear Science*, vol. 50, no. 4, pp. 1103-1110, Aug, 2003.

[72] L. J. Meng, and N. H. Clinthorne, "A modified uniform Cramer-Rao bound for multiple pinhole aperture design," *IEEE Transactions on Medical Imaging*, vol. 23, no. 7, pp. 896-902, Jul, 2004.

[73] L. J. Meng, and N. Li, "A Vector Uniform Cramer-Rao Bound for SPECT System Design," *IEEE Transactions on Nuclear Science,* vol. 56, no. 1, pp. 81-90, Feb, 2009.
[74] L. J. Meng, W. L. Rogers, N. H. Clinthorne *et al.*, "Feasibility study of Compton scattering enhanced multiple pinhole imager for nuclear medicine," *IEEE Transactions on Nuclear Science,* vol. 50, no. 5, pp. 1609-1617, Oct, 2003.
[75] H. C. Gifford, P. E. Kinahan, C. Lartizien *et al.*, "Evaluation of multiclass model observers in PET LROC studies," *IEEE Transactions on Nuclear Science,* vol. 54, no. 1, pp. 116-123, Feb, 2007.
[76] P. Khurd, and G. Gindi, "Fast LROC analysis of Bayesian reconstructed emission tomographic images using model observers," *Physics in Medicine and Biology,* vol. 50, no. 7, pp. 1519-1532, Apr 7, 2005.
[77] W. E. Smith, and H. H. Barrett, "Hotelling Trace Criterion as a Figure of Merit for the Optimization of Imaging-Systems," *Journal of the Optical Society of America a-Optics Image Science and Vision,* vol. 3, no. 5, pp. 717-725, May, 1986.
[78] H. H. Barrett, C. K. Abbey, and E. Clarkson, "Objective assessment of image quality. III. ROC metrics, ideal observers, and likelihood-generating functions," *Journal of the Optical Society of America a-Optics Image Science and Vision,* vol. 15, no. 6, pp. 1520-1535, Jun, 1998.
[79] H. C. Gifford, R. G. Wells, and M. A. King, "A comparison of human observer LROC and numerical observer ROC for tumor detection in SPECT images," *IEEE Transactions on Nuclear Science,* vol. 46, no. 4, pp. 1032-1037, Aug, 1999.
[80] J. Y. Qi, and R. H. Huesman, "Theoretical study of lesion detectability of MAP reconstruction using computer observers," *IEEE Transactions on Medical Imaging,* vol. 20, no. 8, pp. 815-822, Aug, 2001.
[81] J. Qi, and R. H. Huesman, "Lesion detectability of MAP reconstruction using computer observer: A theoretical study.," *Journal of Nuclear Medicine,* vol. 41, no. 5, pp. 18p-19p, May, 2000.
[82] P. Khurd, and G. Gindi, "Rapid computation of LROC figures of merit using numerical observers (for SPECT/PET reconstruction)," *IEEE Transactions on Nuclear Science,* vol. 52, no. 3, pp. 618-626, Jun, 2005.
[83] L. Zhou, P. Khurd, S. Kulkarni *et al.*, "Aperture optimization in emission imaging using ideal observers for joint detection and localization," *Physics in Medicine and Biology,* vol. 53, no. 8, pp. 2019-2034, Apr 21, 2008.
[84] G. Bal, G. L. Zeng, R. M. Lewitt *et al.*, "Study of different pinhole configurations for small animal tumor imaging," in IEEE NSS/MIC, Philadelphia, PA, 2004.
[85] R. Accorsi, F. Gasparini, and R. C. Lanza, "Optimal coded aperture patterns for improved SNR in nuclear medicine imaging," *Nuclear Instruments & Methods in Physics Research Section a-Accelerators Spectrometers Detectors and Associated Equipment,* vol. 474, no. 3, pp. 273-284, Dec 11, 2001.
[86] H. L. V. Trees, *Detection, Estimation, and Modulation Theory*, New York: Wiley, 1968.
[87] C. H. Hua, N. H. Clinthorne, S. J. Wilderman *et al.*, "Quantitative evaluation of information loss for Compton cameras," *IEEE Transactions on Nuclear Science,* vol. 46, no. 3, pp. 587-593, Jun, 1999.
[88] T. J. Kragh, and A. Hero, "Bias-resolution-variance tradeoffs for single pixel estimation tasks using the uniform Cramer-Rao bound," in IEEE NSS/MIC, Lyon, 2000.
[89] T. J. Kragh, and A. O. Hero, III "Bias-resolution-variance tradeoffs for single pixel estimation tasks using the uniform Cramer-Rao bound," in IEEE NSS/MIC, 2000.
[90] J. Nocedal, and S. Wright, *Numerical Optimization*, 2 ed.: Springer, 2006.
[91] Y. C. Eldar, "Minimum variance in biased estimation: Bounds and asymptotically optimal estimators," *IEEE Transactions on Signal Processing,* vol. 52, no. 7, pp. 1915-1930, Jul, 2004.

[92] J. A. Fessler, and S. D. Booth, "Conjugate-gradient preconditioning methods for shift-variant PET image reconstruction," *IEEE Transactions on Image Processing*, vol. 8, no. 5, pp. 688-699, May, 1999.
[93] L. J. Meng, J. W. Tan, and G. Fu, "Design study of an MRI compatible ultra-high resolution SPECT for in vivo mice brain imaging," in IEEE NSS/MIC, 2007.
[94] L. J. Meng, and N. Li, "Non-Uniform Object-Space Pixelation (NUOP) for Penalized Maximum-Likelihood Image Reconstruction for a Single Photon Emission Microscope System," *IEEE Transactions on Nuclear Science*, vol. 56, no. 5, pp. 2777-2788, Oct, 2009.
[95] J. Y. Qi, and R. M. Leahy, "A theoretical study of the contrast recovery and variance of MAP reconstructions from PET data," *IEEE Transactions on Medical Imaging*, vol. 18, no. 4, pp. 293-305, Apr, 1999.
[96] N. Li, and L.-J. Meng, "Adaptive angular sampling approach for emission tomography." in IEEE NSS/MIC, 2010.
[97] S. D. Metzler, J. E. Bowsher, K. L. Greer *et al.*, "Analytic determination of the pinhole collimator's point-spread function and RMS resolution with penetration," *IEEE Transactions on Medical Imaging*, vol. 21, no. 8, pp. 878-887, Aug, 2002.
[98] S. D. Metzler, J. E. Bowsher, M. F. Smith *et al.*, "Analytic determination of pinhole collimator sensitivity with penetration," *IEEE Transactions on Medical Imaging*, vol. 20, no. 8, pp. 730-741, Aug, 2001.
[99] T. Funk, D. L. Kirch, M. S. Sun *et al.*, "Simulation and validation of point spread functions in pinhole SPECT imaging," *IEEE Transactions on Nuclear Science*, vol. 53, no. 5, pp. 2729-2735, Oct, 2006.
[100] J. A. Fessler, "Spatial resolution and noise tradeoffs in pinhole imaging system design: a density estimation approach," *Optics Express*, vol. 2, no. 6, pp. 237-253, Mar 16, 1998.
[101] Z. J. Bai, M. Fahey, and G. Golub, "Some large-scale matrix computation problems," *Journal of Computational and Applied Mathematics*, vol. 74, no. 1-2, pp. 71-89, Nov 5, 1996.
[102] W. H. Press, S. A. Teukolsky, and W. T. Vettering, *Numerical Recipes in C++: The Art of Scientific Computing*, 2nd ed.: Cambridge,U.K.: Cambridge Univ., 2007.
[103] "http://www.mathworks.com/products/matlab/."
[104] L.-J. Meng, J. C. Zhang, Z. M. Shen *et al.*, "Development of Ultra-Compact CdTe and CZT Detectors for Sub-500um Resolution PET Imaging Inside a Small-Bore 14.1 T MRI/MRS Scanner," in The Imaging at Illinois, Champaign, IL, 2012
[105] "User Guide of University of Illinois shared computing cluster," https://campuscluster.illinois.edu/.